

Al-Diban, Sabine; Seel, Norbert M.

## **Evaluation als Forschungsaufgabe von Instruktionsdesign. Dargestellt am Beispiel einer multimedialen Lernumgebung**

*Unterrichtswissenschaft 27 (1999) 1, S. 29-60*



Quellenangabe/ Reference:

Al-Diban, Sabine; Seel, Norbert M.: Evaluation als Forschungsaufgabe von Instruktionsdesign. Dargestellt am Beispiel einer multimedialen Lernumgebung - In: Unterrichtswissenschaft 27 (1999) 1, S. 29-60 - URN: urn:nbn:de:0111-opus-77277 - DOI: 10.25656/01:7727

<https://nbn-resolving.org/urn:nbn:de:0111-opus-77277>

<https://doi.org/10.25656/01:7727>

in Kooperation mit / in cooperation with:

# **BELTZ JUVENTA**

<http://www.juventa.de>

### **Nutzungsbedingungen**

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, veröffentlichen oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

### **Terms of use**

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

### **Kontakt / Contact:**

peDOCS  
DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation  
Informationszentrum (IZ) Bildung  
E-Mail: [pedocs@dipf.de](mailto:pedocs@dipf.de)  
Internet: [www.pedocs.de](http://www.pedocs.de)

Digitalisiert

Mitglied der

  
Leibniz-Gemeinschaft

---

# Unterrichtswissenschaft

Zeitschrift für Lernforschung  
27. Jahrgang / 1999 / Heft 1

---

## Thema: Instruktionsdesign

Verantwortlicher Herausgeber:  
Norbert M. Seel

Norbert Seel: Instruktionsdesign: Modelle und Anwendungen	2
Helmut M. Niegemann, Eva-Maria Gronki-Jost, Oliver Neff: Instruktionsdesign zur Förderung des selbständigen Erwerbs theoretischen Wissens in der kaufmännischen Berufsausbildung	12
Sabine Al-Diban, Norbert M. Seel: Evaluation als Forschungsaufgabe von Instruktionsdesign	29
Günter Dörr: Didaktisches Design multimedialer Lernumgebungen in der betrieblichen Weiterbildung	61
Sanne Dijkstra, Gerialien Holsbrink-Engels: Research on Instruction in Multimedia Learning Environments: Design and Effects	78

---

Sabine Al-Diban, Norbert M. Seel

## **Evaluation als Forschungsaufgabe von Instruktionsdesign**

Dargestellt am Beispiel einer multimedialen Lernumgebung

Evaluation as a Research Problem of Instructional Design

---

*Der Artikel beschreibt anhand eines konkreten Beispiels die verschiedenen Phasen der Evaluation eines multimedialen Lehrprogramms zum Thema „Wirtschaftssysteme“. Zugrunde gelegt wird dabei ein von Ross & Morrison (1997) entwickeltes Evaluationsmodell mit vier Komponenten, von denen zwei – die Methodologie und Datenanalyse – ausführlich in Bezug auf die Evaluation des Multimedia-Programms dargestellt werden, das nach den Prinzipien der Lehrkonzeption „Cognitive Apprenticeship“ entwickelt und gestaltet worden ist. Mit der vorliegenden Untersuchung sind erstmals alle Methoden der „Cognitive Apprenticeship“ (Modeling, Coaching, Scaffolding, Artikulation, Reflexion und Exploration) innerhalb eines Lehrprogramms realisiert und bezüglich ihrer Lerneffektivität untersucht worden. Die Ergebnisse stützen generell die Effektivität der Lehrkonzeption der „Cognitive Apprenticeship“ als Grundlage für die didaktische Gestaltung multimedialer Lehrsysteme. Das Multimedia-Programm wurde aufgrund der Evaluationsdaten einer grundlegenden Modifizierung und Revision unterzogen, um die Umsetzung der „Cognitive Apprenticeship“ auch auf inhaltlicher und methodischer Ebene zu verbessern. Weiterhin bleibt zu untersuchen, wie die recht hohen inter- wie auch intraindividuellen Leistungsunterschiede zustande kommen.*

*The article describes the different phases of evaluation applied to the multimedia program „Economic systems“ as an example. Taking the evaluation model of Ross & Morrison (1997) as a fundamental basis, two of the four components – namely methodology and data analysis – have been applied in order to evaluate the multimedia program that has been developed and designed in accordance to the principles of the „Cognitive Apprenticeship“-approach.*

*For the first time, all methods of the Cognitive Apprenticeship (i.e. modeling, coaching, scaffolding, articulation, reflection and exploration) have been realized within a unique multimedia learning environment, and have been investigated with regard to their learning effectivity. The results generally indicate the effectivity of the Cognitive Apprenticeship methods considered as a theoretically sound conception for the design of multimedia learning environments. In view of the evaluation data the program has been modified and revised in order to improve the learning environment with regard to its contents and methods. In future, we will investigate in more detail how the observed substantial interindividual as well as intraindividual differences in performance could be explained.*

### **1. Einführung in die Problemstellung**

„Instruktionsdesign“ bzw. „didaktisches Design“ bezeichnet eine pädagogisch-psychologische und systemtheoretisch begründete Technologie der

Entwicklung, Gestaltung, Implementation und Evaluation von Lernsituationen und Lernumgebungen – angepaßt an „Lerner (Zielgruppe), Aufgaben (Anforderungen), Ressourcen und Bezugssystem“ (Flehsig, 1990, S. 31). Bezüglich der hierfür erforderlichen Planungsschritte besteht in der einschlägigen Literatur seit langem weitgehender Konsens – so auch im Hinblick auf eine systematische Evaluation der nach didaktischen Gesichtspunkten entwickelten und gestalteten Lernumgebungen und Lernmaterialien. Es dürfte unmittelbar einleuchten, daß eine genaue Prüfung der Effektivität von Lehr-Lern-Systemen vor allem dann bedeutsam wird, wenn diese unter Nutzung kostspieliger Informations- und Kommunikationstechnologien entwickelt und implementiert werden. Darauf hatte bereits Stolurow (1973) bei der Darstellung der transaktionalen Unterweisung und ihrer notwendigen Unterstützung durch Computer hingewiesen, und Seel und Dörr (1997) haben dementsprechend die Bedeutung der Kontrolle der Lerneffektivität multimedialer Lernumgebungen auch unter Kosten-Nutzen-Aspekten hervorgehoben.

Die Evaluation von Lernumgebungen und Lernmaterialien auf eine Analyse der Kosten und Nutzen zu beschränken, entspräche aber nicht dem Stellenwert, der ihr im Instruktionsdesign als Planungswissenschaft insgesamt zugesprochen wird, insofern alle bekannten Modelle des Instruktionsdesigns eine methodisch solide durchzuführende Evaluation als zentralen und unverzichtbaren Bestandteil eines effektiven, rational begründeten didaktischen Designs von Lehr-Lern-Prozessen betrachten (vgl. im Überblick Dijkstra et al., 1997; Tennyson et al., 1997). Dabei wird grundsätzlich zwischen zwei Formen der Evaluation unterschieden: Die *formative Evaluation* zielt darauf, Lehrprogramme zu verbessern, indem Informationen über die effektive Nutzung der verschiedenen Programmkomponenten in das System zurückgemeldet werden. Im Unterschied dazu wird eine *summative Evaluation* durchgeführt, um das Ausmaß zu messen, in dem angezielte Hauptergebnisse am Ende eines Kurses erreicht werden. Eine weitere Unterscheidung ist hinsichtlich der Instrumentierung der Evaluation zu treffen, insofern diese auf die Erhebung quantitativer wie auch qualitativer Daten angelegt sein kann. Zu guter Letzt kann auch danach differenziert werden, ob der Erfolg eines Lehrprogramms durch das Ausmaß definiert wird, in dem die festgestellten Lernergebnisse den festgelegten Zielkriterien entsprechen, oder ob Lernen als ein idiosynkratischer Prozeß der Wissenskonstruktion begriffen wird, dessen bedeutsame Ergebnisse in Einstellungen, Wertvorstellungen, einem Gefühl von Verantwortlichkeit oder den mentalen Modellen bestehen, die Personen während ihrer Lernerfahrungen entwickeln.

Bereits diese knappe Skizze der Zielsetzungen und Aufgabenstellungen der Evaluation verdeutlicht, daß diese Komponenten des Instruktionsdesigns nicht nur eine aufwendige praktische Tätigkeit ist, sondern eine relativ umfassende und komplexe Forschungsaufgabe darstellt, der in großen Entwicklungsprojekten wie z. B. dem Jasper-Projekt der Cognition and Technology Group at Vanderbilt (1997), dem Projekt „Science Quest“ (Litchfield & Mattson, 1989) oder den „Goal-based Scenarios“ (z. B. Schank et al. 1993/94) zentrale Bedeutung beigemessen wird. Auch im Rahmen des „4M-Projek-



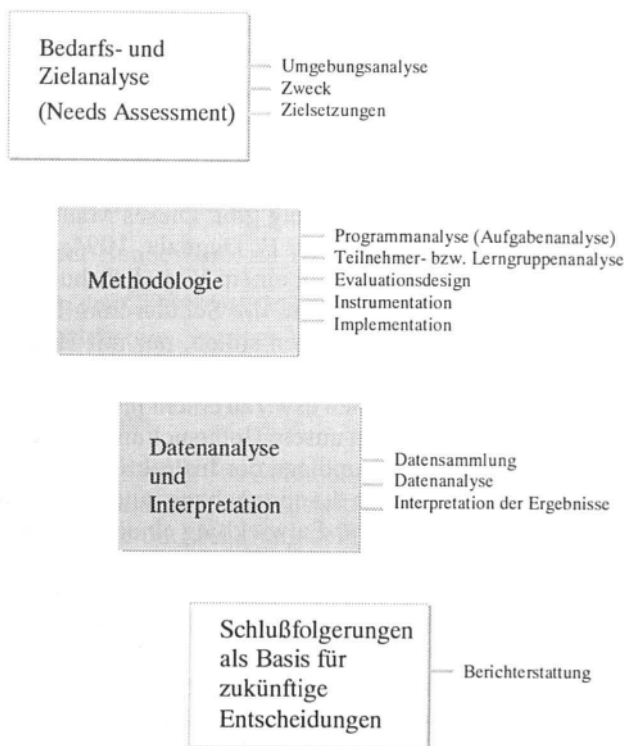
tes“ (*Multi-Media und Mentale Modelle*)<sup>1</sup>, das sich mit der Konstruktion mentaler Analogiemodelle unter den besonderen Bedingungen des Lernens mit multimedialen Lehrsystemen befaßt (vgl. Al-Diban et al., i.pr.; Seel et al., 1998), stellte sich die Evaluation des nach den didaktischen Prinzipien der Konzeption „Cognitive Apprenticeship“ entwickelten und gestalteten Multimedia-Programms „Wirtschaftssysteme“ als zentrale Untersuchungsaufgabe heraus. Darüber soll im folgenden berichtet werden.

## 2. Das Evaluationsmodell

Unter den verschiedenen Ansätzen, die in der Literatur für die Evaluation multimedialer Lehrprogramme beschrieben werden, entschieden wir uns für das Evaluationsmodell von Ross und Morrison (1997), dessen Hauptkomponenten (nach geringfügiger Modifikation) wie folgt darzustellen sind:

Abbildung 1:

Hauptkomponenten des Evaluationsmodells von Ross & Morrison (1997, S. 340)



<sup>1</sup> Dieses Projekt wird seit 1994 von der Deutschen Forschungsgemeinschaft (Se 399/4) gefördert. Außer den Autoren gehören der Projektgruppe an: Frau Dipl.-Psych. S. Held, Frau Dipl. Inform. C. Hess, Frau Katharina Schenk und Frau Dipl.-Psych. S. Wilczek.

An diesem Modell orientierte sich die Evaluation des Multimedia-Programms „Wirtschaftssysteme“, über die wir nun berichten werden. Wie in der Abbildung hervorgehoben, werden wir uns dabei vor allem auf die Komponenten „*Methodologie*“ und „*Datenanalyse und Interpretation*“ konzentrieren. Dementsprechend stehen folgende Schritte der Programmevaluation im Mittelpunkt unserer Betrachtungen: (1) Die *Programmanalyse*, um die Inhalte und Vermittlungsmethoden des Lehrprogramms zu bestimmen, (2) eine *Teilnehmer-Analyse*, um die Lerngruppe(n) und den Anwendungsbeereich des Programms festzulegen, (3) eine *Spezifikation des Evaluationsdesigns*, (4) die Entwicklung bzw. *Festlegung der (Meß-) Instrumente* und (5) die *Ausführung des Evaluationsplans* und Kontrolle der Ausführung. Daran schließt die *Datenanalyse und Interpretation* der Ergebnisse an, gefolgt von Schlußfolgerungen hinsichtlich evtl. notwendiger Modifikationen des Lehrprogramms.

Abgesehen davon, daß die Evaluation des Lehrprogramms „Wirtschaftssysteme“ im Rahmen des „4M-Projektes“ unerläßlich war, um zu fundierten Aussagen hinsichtlich seiner Lern- und Transferwirksamkeit zu gelangen, begreifen wir sie auch als Beitrag zur empirischen Begründung einer Multimedia-Didaktik, in deren Entwicklung wir noch stehen. Tatsächlich ist diesbezüglich kritisch festzuhalten, daß es zur Zeit (noch) keine umfassende und einheitliche Multimedia-Didaktik im Sinne eines theoriebasierten didaktischen Designs multimedialer Lernsysteme gibt — so das Urteil von Issing (1994) und Goodyear (1997). Vielmehr gibt es so viele erfolversprechende Ansätze des didaktischen Designs multimedialer Lernumgebungen wie es Institute und Zentren für ihre Entwicklung gibt. Dieses Manko springt um so stärker ins Auge, als viele Autoren (z. B. Dugdale, 1994; Hannafin, 1992; Rieber, 1996), wie bereits Stolurow vor einem Vierteljahrhundert, davon ausgehen, daß effektive Lernumgebungen, die Schüler zum Nachdenken und Problemlösen anregen und qualifizieren sollen, nur mit Hilfe „emergenter Technologien“ zu realisieren seien, die verschiedene Medienkomponenten (Text, Sprache, Video, Simulationen usw.) zu einem integrativen Verbund zusammenschweißen. Insofern stellt unsere Untersuchung der Lern- und Transferwirksamkeit einen auf der Grundlage der Instruktionsstrategie „Cognitive Apprenticeship“ entworfenen Lernumgebung und ihrer multimedialen Ausgestaltung auch ein Beitrag zur Entwicklung einer Multimedia-Didaktik dar und ist im Kontext mit anderen Forschungsansätzen wie z. B. dem „Jasper-Projekt“ der Cognition and Technology Group at Vanderbilt (1997) und den „Goal-based Scenarios“ von Schank und Mitarbeitern (z. B. Schank et al. 1993/94) zu sehen.

### **3. Methodologie der Evaluation des Lehrprogramms „Wirtschaftssysteme“**

Gemäß dem zugrundegelegten Evaluationsmodell besteht der erste Schritt der Hauptkomponente „Methodologie“ in der Programmanalyse, die sich auf Inhalt und Methode der Vermittlung bezieht.

### 3.1 Programmanalyse.

In bezug auf die *Inhalte* orientierten wir uns grundlegend an einer Aussage Glasers (1984), wonach eine Verbesserung von Lernfähigkeiten durch die Einübung begrifflichen und prozeduralen Wissens im Kontext spezifischer Wissensbereiche erfolgt. Lernen und Denken entwickeln sich in weitgehender Abhängigkeit von den Inhalten und Begriffen eines Wissensbereichs, wie sie in Lernsituationen enthalten sind. Diese schränken das Wissen ein, um bestimmten Zwecken und Zielen zu dienen. Den thematischen Schwerpunkt des Lehrprogramms stellen *Wirtschaftssysteme als erdachte Gebilde* dar, mittels derer reale Wirtschaftsordnungen erklärt werden. Damit entspricht das Programm aus inhaltlicher Sicht der zentralen Bedeutung, die volkswirtschaftliche Prozesse und Entscheidungen für unser alltägliches Leben haben: „All your life - from cradle to grave and beyond — you will run against the brute truths of economics“ (Samuelson, 1983, p. ii). Diese Annahme liegt auch den meisten Lehrplänen z. B. für die gymnasiale Oberstufe zugrunde, insofern die Behandlung volkswirtschaftlicher Themen verbindlich vorgeschrieben wird. Ihnen entsprechend thematisiert das von uns entwickelte Multimedia-Programm dynamische Wirtschaftssysteme anhand von Wirtschaftskreisläufen und der Geldpolitik der Bundesbank. Im einzelnen werden — in Entsprechung mit Glasers Verdikt — folgende Themen behandelt:

- die Unterscheidung der verschiedenen Wirtschaftssektoren („private Haushalte“, „Unternehmen“, „Staat“, „Finanzsektor“ und „Ausland“), inklusive der zwischen ihnen bestehenden Geldströme,
- die Stellung der Bundesbank im Wirtschaftssystem der BRD,
- die Festlegung der Geldmenge M3 als Grundlage für die Feststellung des Geldmengenwachstums in einer Volkswirtschaft,
- die Möglichkeiten der Geldschöpfung durch die Bundesbank und durch die Geschäftsbanken,
- die mit der Überwachung der Geldmenge M3 verfolgten geldpolitischen Ziele der Bundesbank, d. h. die ausreichende Versorgung der Volkswirtschaft mit Geld und die Sicherung der Stabilität des Geldwertes.

Die Lehrinhalte werden mit dem Ziel, die Schüler zur Konstruktion adäquater Analogiemodelle dynamischer Wirtschaftssysteme zu befähigen, unter konsequenter Nutzung von „konzeptuellen Modellen“ vermittelt (Seel, 1991). Als geeignetes konzeptuelles Modell des Wirtschaftssystems wird — wie in fast allen einschlägigen Lehrbüchern — das Kreislaufmodell eingesetzt (vgl. weiterführend: Seel et al., 1998).

Der *Vermittlung* der Lehrinhalte wird die Lehrkonzeption der „Cognitive Apprenticeship“ zugrundegelegt. Sie ragt unter den neueren Entwürfen „kognitiver Lehr-Lern-Theorien“ deshalb heraus, da sie darauf zielt, Merkmale „idealer“ Lernumgebungen zu identifizieren und in ein Gesamtmodell effek-

tiven Lehrens und Lernens zu integrieren. In Anlehnung an Verfahren der Meta-Analyse gelangten die Autoren der Cognitive Apprenticeship (Collins, Brown & Newman 1989; Collins, Brown & Holum, 1993) zu insgesamt 18 Merkmalsdimensionen von Lernumgebungen, die sie in vier umfassende Kategorien gruppierten: Inhalte, Methoden, Sequenzierung und Soziologie. Die ausführliche Beschreibung dieser Kategorien in bezug auf das Lehrprogramm „Wirtschaftssystem“ findet sich in Seel et al. (1998). Im folgenden konzentrieren wir uns auf die aus der Sicht des zugrundegelegten Evaluationsmodells relevanten Kategorien.

Hinsichtlich der Kategorie *Inhalte* unterscheidet die Konzeption der Cognitive Apprenticeship zwischen den Formen *bereichsspezifisches* (deklaratives) *Wissen*, *heuristisches Wissen*, *Kontrollstrategien* und individuelle *Lernstile*. Das bereichsspezifische Wissen, das die Schüler durch die Arbeit mit dem Lehrprogramm erwerben sollen, hat seinen thematischen Schwerpunkt im Bereich Wirtschaftskreisläufe (insbesondere der Geldpolitik der Bundesbank), was den Lehrplänen für die gymnasiale Oberstufe für das Fach Gemeinschaftskunde/Rechtserziehung/Wirtschaft entspricht. In bezug auf den Erwerb heuristischen Wissens zielt das Lehrprogramm darauf, die Lernenden im Gebrauch der Heuristiken des Zerlegens einer komplexen Aufgabe in einfacher zu lösende Teilaufgaben und der Analogiebildung zu schulen. Durch das im Lehrprogramm implementierte Training handlungsbezogener Metakognition wird auch ein Beitrag zur Entwicklung von Kontrollstrategien geleistet.

In der Kategorie *Methoden* unterscheidet die Konzeption der Cognitive Apprenticeship zwischen den Lehrverfahren des *Modeling* (= Vormachen einer Aufgabenlösung), *Coaching* (= Nachhelfen und Unterstützen beim Nachmachen) und *Scaffolding* (= Entwickeln aufgabenbezogener Heuristiken, die sich an dem Prinzip des darstellenden Erklärens orientieren) sowie den eher auf die Förderung metakognitiven Lernens bezogenen Verfahren *Artikulation* (= Selbstverbalisierung i. S. lauten Denkens), *Reflexion* (= Nachdenken über eigene Lösungsansätze) und *Exploration* (= Erkundung neuer Aufgabenstellungen).

Die starke Betonung darstellenden Erklärens liegt darin begründet, daß sich die Autoren der Cognitive Apprenticeship auf die Metapher der handwerklichen Meisterlehre und Lehrlingsausbildung beziehen und damit dem seit der Antike bekannten Verständnis von Lehren und Lernen als dem gemeinsamen methodischen Handeln zweier oder mehrerer Partner entsprechen, durch das sich wenigstens einer von ihnen Lerngegenstände (hier als praktische oder theoretische Fertigkeiten definiert) aneignet und dabei vermittels geeigneter Handlungen des anderen (z. B. durch Vormachen und Hilfestellungen bei der Ausführung einer Fertigkeit) angeleitet und unterstützt wird. Infolgedessen gründen die ersten drei Methoden der Cognitive Apprenticeship auf der Annahme, daß die Leistungsformen von Lernenden verbessert werden können, wenn sie beobachten, wie Experten mit gestellten Aufgaben umgehen und Probleme lösen. Dem entspricht das Modeling: Die Beobachtung eines Ex-

perten bei der Problemlösung soll bewirken, daß die Lernenden ein Modell des Lösungsprozesses entwerfen und sodann selbst benutzen, um ähnliche Probleme zu lösen. Aus der traditionellen Meisterlehre aber weiß man, daß der Lehrling nicht durch bloßes Zuschauen lernt. Er muß auch versuchen, die Tätigkeiten des Meisters aktiv nachzumachen. Je schwieriger die Fertigkeit auszuführen ist, desto mehr Fehler macht der Lehrling, und der Meister muß korrigierend eingreifen. In Entsprechung dazu zielt Coaching darauf ab, den Lernenden beim Nachahmen der Problemlösung zu beobachten, um ihm Hilfen, Hinweise, Rückmeldungen, Erinnerungshilfen usw. anzubieten, wenn er Fehler macht oder nicht weiterkommt. Dabei kann es notwendig werden, ihm weitergehende Hilfestellungen anzubieten, um den Aufbau eines problemlösungsrelevanten Handlungsgerüsts zu unterstützen. Dies wird als *Scaffolding* bezeichnet. Die Zerlegung eines Ziels in Teilziele und die Entwicklung von Strukturdiagrammen zur graphischen Darstellung des Problemraums gehören ebenso dazu wie die Bereitstellung physikalischer Modelle, um die Problemstellung zu veranschaulichen. Mit zunehmender Kompetenz des Lernenden, die Aufgaben zu lösen, werden die Hilfestellungen ausgeblendet.

Schüler dazu zu bringen, sich ihr Wissen und die Versuche, ein Problem zu lösen, bewußt zu machen, steht im Mittelpunkt der auf metakognitive Kontrolle („Monitoring“ und „Selbstregulation“) des Lernens zielenden Methoden. Hier nennen die Autoren der Cognitive Apprenticeship zunächst die *Artikulation* als Verbalisierung eigener Gedanken beim Problemlösen, was durchaus sinnvoll ist, belegt doch eine Vielzahl von Untersuchungen, daß „lautes Denken“ oft hilfreich ist, um ein Problem zu lösen. In einem nächsten Schritt geht es darum, die eigenen Problemlösungsversuche systematisch und selbstkritisch mit denen des Experten zu vergleichen. Das wird als *Reflexion* bezeichnet. Den Schlußpunkt setzt die *Exploration*: Die Lernenden sollen bei neuen Aufgaben eigene Lösungshypothesen und -wege entwickeln und die erworbenen Lernstrategien anwenden.

Über die bisher beschriebenen Merkmalsdimensionen der Cognitive Apprenticeship hinausgehend orientiert sich das Lehrprogramm „Wirtschaftssysteme“ auch explizit an der Kategorie *Sequenzierung* der Lernaufgaben und legt hierfür die Prinzipien der zunehmenden Komplexität, der zunehmenden Vielfalt und der Behandlung allgemeiner vor spezifischen Fertigkeiten zugrunde. Das erste Prinzip kann auch als „vom Einfachen zum Schwierigen“ umschrieben werden, indem zunächst relativ einfache Probleme eingeführt werden, deren Schwierigkeitsgrad dann schrittweise gesteigert wird. Danach sollen Probleme eingeführt werden, die verschiedenartige Strategien oder Fertigkeiten zu ihrer Lösung erfordern („zunehmende Vielfalt“), um die Übertragung des Gelernten auf neue Aufgaben zu unterstützen. Als förderlich wird dabei das Vorgehen beurteilt, eher allgemeine als spezielle Fertigkeiten der Problemlösung zu trainieren.

Bevor wir in der Darstellung der Evaluationsschritte fortfahren, sei der Hinweis erlaubt, daß im Verhältnis zu der Menge der theoretischen Abhandlun-

gen zur Cognitive Apprenticeship vergleichsweise wenig praktische Realisierungen und Evaluationen vorliegen. Dies ist in Anbetracht der Bedeutung, die ihr mittlerweile als Lehrstrategie zugeschrieben wird (vgl. Choi & Hannafin, 1995), bemerkenswert. Die Analyse der wenigen empirischen Untersuchungen läßt die Schlußfolgerung zu, daß die von Collins et al. (1989) genannten Merkmale von Lernumgebungen vielfach als „Bausteine“ betrachtet werden, die einzeln und voneinander getrennt eingesetzt und hinsichtlich ihrer Lerneffektivität geprüft werden können. Tatsächlich konzentrierte man sich bisher vor allem auf die Untersuchung von *Modeling* und *Coaching* (vgl. Farmer, Buckmaster & Legrand, 1992; Lajoie & Lesgold, 1989; Niegemann, 1995; Volet, 1991); andere Studien (z. B. Järvelä, 1995) untersuchten die Methoden *Modeling*, *Scaffolding* und *Reflection* im Verlaufe verbaler Interaktionen im Unterricht. Nur eine Untersuchung (Casey, 1996) ist uns bisher bekannt geworden, in der versucht wurde, alle Methoden der Cognitive Apprenticeship hinsichtlich ihrer Lerneffektivität zu evaluieren. Diese Untersuchung ist für uns auch deshalb interessant, da sie der Frage nachging, ob und in welchem Umfange die Merkmale der Cognitive Apprenticeship in ein multimediales Lernsystem inkorporierbar sind.

### 3.2 Teilnehmer- bzw. Lerngruppenanalyse

Das Multimedia-Programm „Wirtschaftssysteme“ wendet sich primär an Schüler/-innen der Oberstufe des Gymnasiums mit nichtwirtschaftlicher Ausrichtung (vorzugsweise Klassenstufe 12/I), kann aber auch, wie unsere ersten Evaluationsstudien belegen, erfolgreich von Studierenden (nichtwirtschaftlicher Disziplinen) zum selbstkontrollierten Lernen benutzt werden.

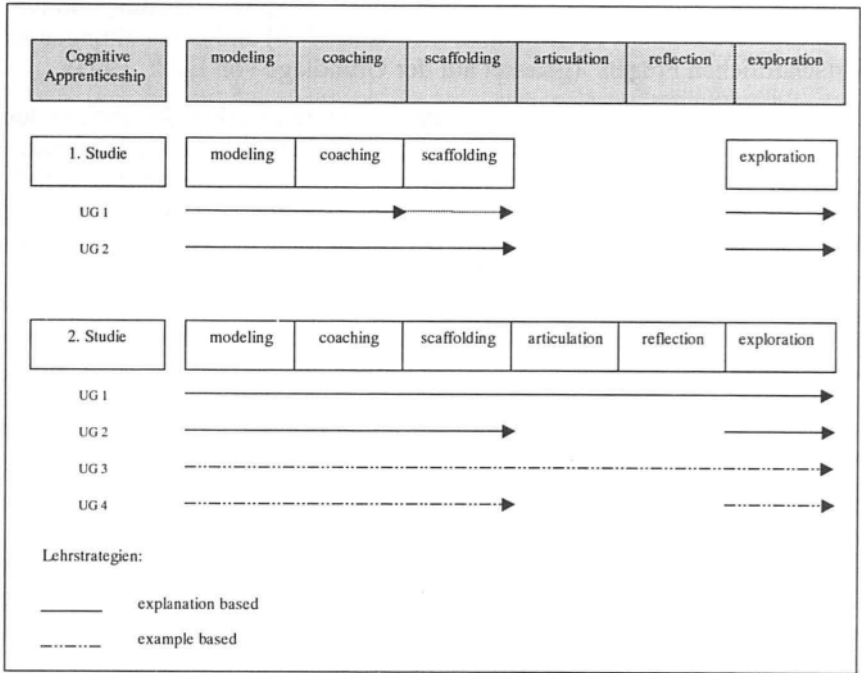
Das Lehrprogramm bzw. einzelne Teile daraus wurden bislang in vier Untersuchungen einer Evaluation unterzogen. Die erste Untersuchung befaßte sich mit der Evaluation der Methode *Modeling* und kann aus verschiedenen Gründen als eine explorierende *Vorstudie* verstanden werden. Im folgenden berichten wir über die erste und zweite Studie; die dritte Untersuchung wird zur Zeit durchgeführt. An Studie 1 nahmen Studierende der Technischen Universität Dresden (N = 36, Lehramtsstudenten und Studierende im Hauptfach Erziehungswissenschaft) teil, von denen 34 in die Auswertung einbezogen werden konnten. 29 der 34 Versuchspersonen waren weiblich, fünf männlichen Geschlechts, das Durchschnittsalter betrug 21 Jahre. An der 2. Studie nahmen Gymnasiasten der 12. Klassen (N = 84) aus fünf Gymnasien allgemeinbildender Ausrichtungen teil.

### 3.3 Evaluationsdesign

In Übereinstimmung mit der Leitidee pädagogischer Lehr-Lern-Forschung, kumulatives und überdauerndes Lernen komplexer Themenbereiche zu untersuchen (vgl. Norman, 1978), ist das Lehrprogramm „Wirtschaftssysteme“ zur Zeit auf eine Bearbeitungszeit von mehr als sechs Stunden angelegt (das entspricht acht Unterrichtsstunden).

In Anbetracht des Forschungsstandes, der Fragestellung und des innerhalb einer solchen Studie gegebenen Rahmens entschieden wir uns für eine *Strategie der abgestuften formativen Evaluation*. Sie kann wie folgt dargestellt werden:

Abbildung 2: Darstellung der Evaluationsstrategie



Nachdem in der Vorstudie die Effektivität der auf darstellendes Erklären bezogenen Programmteile (*Modeling*, *Coaching*) untersucht worden war, wobei sich die besonders untersuchten kognitiven Strategien „Subsumtion analoger Beispiele unter ein Schema“ (= erklärungs begründetes Lernen) und „Induktion eines Schemas aus analogen Beispielen“ (= ähnlichkeits begründetes Lernen) als gleich wirksam erwiesen hatten, wurden in Studie 1 die spezifischen Lerneffekte von *Scaffolding* nach der erfolgreichen Bearbeitung von *Modeling* und *Coaching* (auf die Leistungen in *Exploration*) geprüft. In Studie 2 stand die formative Evaluation der metakognitiven Komponenten *Artikulation* und *Reflexion* sowie die summative Evaluation des gesamten Lehrprogramms im Zentrum des Forschungsinteresses.

### 3.4 Meßinstrumente

In Anbetracht der zentralen Bedeutung, die dem Vorwissen nach kognitionspsychologischer Auffassung für den Erwerb neuen Wissens zukommt, wurde das Vorwissen der Untersuchungsteilnehmer erstens anhand der Leistungen einer „Einstiegsaufgabe“ zum Verständnis einer Zeitungsmeldung und zweitens mit einem Wissenstest zu Dimensionen der ökonomischen Bildung

(Beck, 1993) ermittelt, der deklaratives Wissen zu den Bereichen Mikro- und Makroökonomie prüft. Um mögliche Interaktionen der experimentellen Bedingungen mit bedeutsamen Persönlichkeitsmerkmalen zu erfassen, wurden verschiedene *Personvariablen* mittels geeigneter Testverfahren erhoben (vgl. dazu im einzelnen Seel et al., 1998). Als relevant erwiesen sich lediglich die Personvariablen „Motivation“ (getestet mittels Leistungsmotivationstest von Hermans, Petermann & Zielinski, 1978) und „Einstellungen zu wirtschaftlichen Fragen“ (getestet auf der Grundlage von Beck, 1993).

Um die Lerneffektivität der verschiedenen experimentellen Variationen des Lehrprogramms zu messen, wurde zunächst die Performanz der Lernenden bei allen Lernaufgaben der verschiedenen Methodenkomponenten verglichen. Auf den jeweiligen Aufgabentypus zur Prüfung der Lernwirksamkeit der verschiedenen Methoden wird in der Darstellung der Befunde noch näher eingegangen. Ein weiteres Maß für die aufgabenspezifische Lernleistung stellten die Fehler der Lernenden sowie die Bearbeitungszeiten pro Lernaufgabe dar. Daneben nahmen die Untersuchungsteilnehmer subjektive Schwierigkeitseinschätzungen der Lernaufgaben vor.

### 3.5 Ausführung des Evaluationsplans

Die Evaluationsuntersuchung mit Schülern wurde mit dem Schulunterricht so abgestimmt, daß das Thema Geldpolitik im Gemeinschaftskundeunterricht noch nicht behandelt worden war. Zur Zeit findet eine weitere Evaluationsuntersuchung statt (wieder mit Gymnasiasten der 12. Klasse), in der die empirische Prüfung einer aufgrund der bisherigen Erkenntnisse modifizierten Umsetzung des didaktischen Konzeptes der Cognitive Apprenticeship durchgeführt wird.

Es folgt eine knappe Skizze<sup>2</sup> der wesentlichen Merkmale des Lehrprogramms „Wirtschaftssysteme“. Der Komponente *Modeling* liegt die von Ahn, Brewer & Mooney (1992) vorgenommene Differenzierung zwischen „erklärungsbegründetem“ und „ähnlichkeitsbegründetem Lernen“ zugrunde.

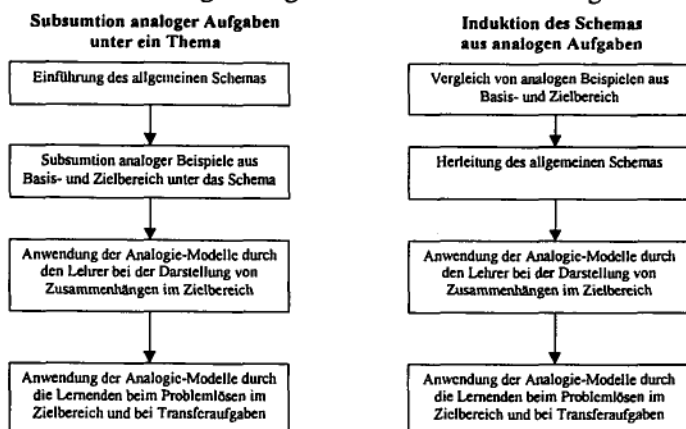
Bezüglich der Gestaltung von *Coaching* und *Scaffolding* folgten wir zum einen der Argumentation von Bransford et al. (1990), wonach effektive „Coaches“ die Leistungen von Lernenden beobachten, um Abweichungen von Zielsetzungen möglichst klein zu halten, zugleich aber so viel Spielraum wie möglich gewähren, um Explorationen und Problemlösen nicht zu unterbinden. Sie helfen den Lernenden, die eigenen Leistungen zu überdenken und mit anderen zu vergleichen. Schließlich zielen effektive Coaches darauf, das Vorwissen der Lernenden zu erfassen und Fehlkonzeptionen sowie fehlerhafte Lösungsstrategien festzustellen, um dies zum Ausgangspunkt konzeptueller Veränderungen zu machen. Zum anderen folgten wir Riedel (1973), der in einer Untersuchung zur Förderung entdeckenden Lernens ergebnis- und

---

<sup>2</sup> Zur ausführenden Darstellung des Lehrprogramms „Wirtschaftssysteme“ ist auf Seel et al. 1998 verwiesen.



Abbildung 3: Kognitive Struktur von Modeling



prozeßorientierte Lernhilfen für die Bearbeitung von Problemstellungen einsetzte, und statteten die *Coaching*-Komponente unseres Lehrprogramms mit analogen Lernaufgaben aus, deren Lösung in Entsprechung mit der Expertenlösung in *Modeling* durch ergebnisorientierte Hilfestellungen angeleitet und unterstützt wurde. Gemäß der Annahme, daß *Scaffolding* zur Herausbildung eines allgemeinen Handlungsgerüsts für Problemlösungen beiträgt, wurden die Lernaufgaben der *Scaffolding*-Komponente mit *prozeßorientierten Hilfestellungen* ausgestattet, die Heuristiken wie die Zerlegung eines Zieles in Teilziele oder Analogiebildung vermitteln sollten (vgl. Catrambone, 1995). Mit fortschreitender Praxis wurden diese Hilfestellungen zurückgenommen („fading“), um eine zunehmende Selbstkontrolle der Aufgabenbearbeitung zu bewirken.

Zusammenfassend kann der Unterschied zwischen *Coaching* und *Scaffolding* wie folgt festgehalten werden:

Abbildung 4: Art der Hilfestellungen bei Coaching und Scaffolding des Lehrprogramms „Wirtschaftssysteme“

COACHING Ergebnisorientierte Hilfen	SCAFFOLDING Problemorientierte Hilfen
<i>Gegenstandszentrierte Organisation der Lernprozesse</i>	<i>Problemzentrierte Organisation der Lernprozesse</i>
<i>Lösungshilfen</i>	<i>Strukturierungshilfen</i>
Hilfen in Form von Aufgabenstellungen und Handlungsanweisungen ohne Bezug auf die Problemstellung	Hilfen zur Präzisierung des Problems
Hilfen zum Identifizieren bedeutsamer Komponenten	Hilfen zum Analysieren des Problems
Hilfen zum Generalisieren des relevanten Beziehungszusammenhangs	Hilfen zur Hypothesenbildung
Hilfen zum Festigen des Gelernten	Hilfen zum Überprüfen der Hypothese
	Hilfen zum Zusammenfassen der Erkenntnisse

*Artikulation* wurde mit dem Ziel eines Effektes auf die Ausbildung von Metakognition durch die Aufforderung des Lehrprogrammes verwirklicht, die Aufgabenstellungen zu paraphrasieren und dazu angefertigte schriftliche Notizen mit einer vom Lernsystem gegebenen Lösung besonders hinsichtlich der eingeführten konzeptuellen Modelle zu vergleichen. Die Methode *Reflexion*, die mit Bezug auf die Expertenlösung darauf zielt zu lehren, wie Probleme eines bestimmten Typs zu lösen sind, wurde in Form einer sog. Backward-Strategie realisiert, um die Aufmerksamkeit der Lernenden auf die allen Problemaufgaben gemeinsame Lösungsstrategie zu lenken. In den Mittelpunkt von *Reflexion* stellten wir daher das Ziel, konzeptuelle Modelle und Analogien in ihrer heuristischen Funktion als Problemlösungsprinzipien zu vermitteln. Während den Lernenden in diesem Programmteil noch Rückmeldungen über eigene Lösungsansätze gegeben werden, zielt *Exploration* darauf, die Lernenden zu ermutigen, verschiedene Strategien der Problemlösung ohne Hilfestellungen auszuprobieren und die Effekte dieser Strategien zu beobachten. Bei der Bearbeitung von insgesamt sieben Transferaufgaben konnten die Lernenden alle verfügbaren Komponenten des Programms (konzeptuelle Modelle, Animationen und Veranschaulichungen, Tips, Bedienungshinweise usw.) benutzen, jedoch ohne weitere problembezogene Hilfen.

## 4. Datenanalyse und Interpretation

Auf dem Hintergrund des skizzierten Evaluationsmodells werden nun einige ausgewählte Daten zur Aufgabenanalyse sowie zur formativen und summativen Evaluation des Lehrprogramms „Wirtschaftssysteme“ dargestellt.

### 4.1 Aufgabenanalyse

Der Ausgangspunkt für die Konstruktion und Sequenzierung einzelner Lernaufgaben im Instruktionsdesign liegt in einer „kognitiven Aufgabenanalyse“ (vgl. Seel, 1997), die sowohl eine semantische Analyse der Aufgabeninhalte (was die Lernenden nach erfolgreichem Unterricht wissen oder können sollen) als auch eine Analyse der für das erfolgreiche Bearbeiten als notwendig und hinreichend erachteten kognitiven Operationen umfaßt. Kognitive Aufgabenanalysen zielen auf die Feststellung der idealen Leistungsformen, gegen die dann die Leistungsformen des einzelnen Lernenden geprüft werden können (vgl. Scandura, 1995).

#### *Objektive Aufgabenanalyse*

Für die zahlreichen und unterschiedlichen Verfahren der Aufgabenanalyse (vgl. Jonassen, Hannum & Tessmer, 1989) gelten verschiedene Kriterien: (1) Für eine hinreichend genaue Differenzierung der Anforderungen und somit Klassifikation von Problemen müssen deren Unterschiede eindeutig und reproduzierbar feststellbar sein; (2) die Aufgabenanalyse sollte unabhängig von einer speziellen Aufgabenstellung branchen- und problemübergreifende Gültigkeit besitzen und (3) unabhängig vom Anwender sein, d. h. keine in-

dividuellen Kenntnisse und Eigenschaften der analysierenden Person enthalten (vgl. Langner, 1991).

Wir verwendeten das Verfahren der Problem- und Aufgabenanalyse von Hacker, Sachse & Schroda (1998), das für komplexe, praxisnahe Aufgabenstellungen entwickelt wurde und in der Lage ist, mehrere Dimensionen voneinander getrennt zu beschreiben. Mit diesem Verfahren kann nicht nur die Komplexität (Anzahl von Teilfunktionen, Anzahl und Stärke ihrer Verknüpfungen) von Lernaufgaben untersucht werden, sondern auch verschiedene Dimensionen wie z. B. „widersprüchliche Ziele“ (Zahl der Ziele, Zahl widersprüchlicher Ziele, Grad der Zielkonkurrenz), „Intransparenz“ (Informationen zu Ausgangsbedingungen, Lösungsweg und angestrebtem Ergebnis), „Freiheitsgrade“ (Lösungsvarianten, Lösungswege), „Dynamik“ (zeitliche Veränderungen von Ausgangsbedingungen, Kalkulierbarkeit von Eingriffen, Einflüsse von Randbedingungen) und „erforderliches Wissen“ (spezifisches Sachwissen, problemspezifische Vorgehensweisen, allgemeine Lösungsstrategien). Jede Lernaufgabe kann somit durch ein multidimensionales, objektives Anforderungsprofil charakterisiert werden.

Vor Durchführung der zweiten Studie wurden die fünf Coachingaufgaben, acht Scaffolding- und sieben Explorationsaufgaben von drei Fachleuten (Absolventen der Betriebswirtschaft/Schwerpunkt Volkswirtschaft) mit dem Fragebogen zur Problem- und Aufgabenanalyse nach Hacker et al. (1998) unabhängig voneinander analysiert. Zweck dieser Analyse war es, zu einer Elimination von Lernaufgaben zu gelangen, da sich bei einem ersten Versuch mit Gymnasiasten eine mehr als siebenstündige Dauer für die Programmbearbeitung herausgestellt hatte, was wir aus lern- und motivationspsychologischen wie auch versuchsorganisatorischen Gründen als problematisch beurteilten. Aufgrund dieser Aufgabenanalyse wurden drei Scaffolding- und zwei Explorationsaufgaben gestrichen, da sie inhaltliche Schwächen und zu hohe Anforderungen an das notwendige Wissen, den Umgang mit Dynamik bzw. eine zu geringe Komplexität und Transparenz aufwiesen.

Die Urteile der Experten hinsichtlich der einzelnen Dimensionen der Aufgabenanalyse erreichten eine mittlere Übereinstimmung von wenigstens 65 % und mehr, was wesentlich zur Erfüllung des Gütekriteriums der Reliabilität der Einschätzungen beiträgt.

Tabelle 1: Urteilerübereinstimmung und Einschätzung der Bedeutsamkeit der Dimensionen in der objektiven Aufgabenanalyse

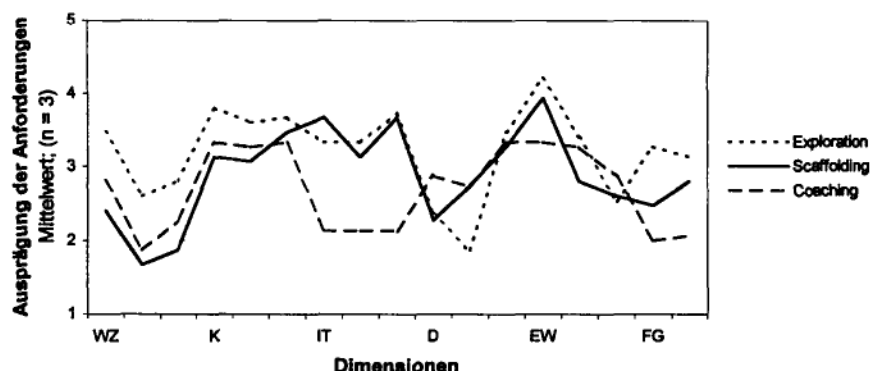
Dimensionen	Wider-sprüch-l. Ziele	Kom-plexi-tät	Intrans-parenz	Freiheits-grade	Dynamik	erforderl. Wissen
mittl. Urteiler-übereinstim-mung [%]	69,8	73,0	72,0	73,0	65,0	78,9
mittl. Bedeutsamkeit der Dimension	1,7	4,3	3,3	2,3	4,0	2,6

Anmerkung: N = 3 unabhängige Experten als Beurteiler; Bedeutsamkeit wurde nach Ratings zwischen 1 (unwichtig) und 5 (sehr wichtig) eingestuft.

Tabelle 1 verdeutlicht, daß die drei Experten übereinstimmend die Dimensionen „Komplexität“, „Dynamik“ und „Intransparenz“ der volkswirtschaftlichen Aufgabenstellungen als am bedeutsamsten für die resultierende Aufgabenschwierigkeit einschätzen, während sie die Dimensionen „Widersprüchliche Ziele“ und „Freiheitsgrade“ als weniger relevant erachteten, was vermutlich mit dem vorrangigen Ziel der Wissensvermittlung im Zusammenhang steht. Bei zwei der wichtigen Dimensionen, „Komplexität“ und „Intransparenz“, wurde eine Übereinstimmung von mehr als 70 % erreicht, die geringere Übereinstimmung für die Dimension „Dynamik“ zeigt die Problematik, die Dynamik volkswirtschaftlicher Aufgabenstellungen einzustufen.

Wie im Ansatz der Cognitive Apprenticeship gefordert, war für die Aufgabengruppen der Komponenten *Coaching*, *Scaffolding* bis zur *Exploration* infolge des zunehmenden Ausmaßes an Komplexität (*Sequenzierung*) und Selbstkontrolle im Lernprozeß (*Fading*) auch ein zunehmender kognitiver Anforderungsgehalt der Lernaufgaben von „leicht“ (im *Coaching*-Abschnitt über „mittel“ im *Scaffolding*-Abschnitt bis „schwer“ im Abschnitt *Exploration*) in den verschiedenen Dimensionen zu erwarten. Tatsächlich aber resultierten für die verschiedenen Aufgabengruppen (in *Coaching*, *Scaffolding* und *Exploration*) nur geringe absolute Unterschiede. Mit anderen Worten: Die Anforderungen aller Aufgaben in den verschiedenen Dimensionen des Fragebogens zur Aufgabenanalyse von Hacker et al. (1998) wurden auf einer Ratingskala zwischen 1 (sehr wenig/sehr schwach) bis 5 (sehr viele/sehr stark) beurteilt und bewegten sich dabei auf einem mittleren Niveau.

Abbildung 5:  
Mittlerer kognitiver Anforderungsgehalt; Gruppen zu je fünf Coaching-, fünf Scaffolding- und fünf Explorationsaufgaben



Eine graduelle Abstufung des Anforderungsgehaltes im Sinne der Cognitive Apprenticeship wurde bei einzelnen Items in den Dimensionen „Komplexität“, „Intransparenz“, „Erforderliches Wissen“ und „Freiheitsgrade“ erreicht (vgl. Tabelle 2).

Tabelle 2:

Mittlere Itemausprägungen der Items zur Objektiven Aufgabenschwierigkeit in Gruppen zu je 5 Coaching-, 5 Scaffolding- und 5 Explorationsfragen

	WZ	WZ	WZ	K	K	K	IT	IT	IT	D	D	D	EW	EW	EW	FG	FG
Explorationsaufgaben . . . . .																	
MW	3,47	2,60	2,80	3,80	3,60	3,67	3,33	3,33	3,73	2,36	1,87	3,47	4,20	3,40	2,53	3,27	3,13
Scaffoldingaufgaben _____																	
MW	2,40	1,67	1,87	3,13	3,07	3,47	3,67	3,13	3,67	2,27	2,73	3,27	3,93	2,80	2,60	2,47	2,80
Coachingaufgaben -----																	
MW	2,80	1,87	2,27	3,33	3,27	3,33	2,13	2,13	2,13	2,87	2,73	3,33	3,33	3,27	2,87	2,00	2,07

Anmerkung: WZ...Widersprüchliche Ziele, K...Komplexität, D...Dynamik, EW...Erforderliches Wissen, FG...Freiheitsgrade.

Die *Komplexität* der Lernaufgaben wird hier von den Experten als besonders bedeutsam für die resultierende Aufgabenschwierigkeit eingestuft. Bei mittleren Anforderungsausprägungen gab es nur geringe Unterschiede zwischen den Aufgabengruppen, wobei die Scaffoldingaufgaben geringfügig „weniger Teilfunktionen und weniger Verknüpfungen“ (Item 1, 2) mit etwas „stärker verknüpften Teilfunktionen“ (Item 3) aufweisen als die Coachingaufgaben. Die Explorationsaufgaben beinhalten bei allen drei Komplexitätsitems die vergleichsweise höchsten Anforderungen. Im Falle der Dimension „Dynamik“ wird die Sequenzierung teilweise „auf den Kopf“ gestellt. Dieses Ergebnis beruht vermutlich auf der Problematik der auch für Experten schwierigen Aufgabe, die Dynamik volkswirtschaftlicher Problemstellungen angemessen einzuschätzen. Bei allgemein „mittlerer“ Dynamik wurden die Explorationsaufgaben als „eher genau kalkulierbar“ eingeschätzt, während Coaching- und Scaffoldingaufgaben sich als nur „mittelmäßig kalkulierbar“ (Item 2) erwiesen. Hinsichtlich der Dimension „Intransparenz“ ist festzustellen, daß die Scaffolding- und Explorationsaufgaben auf einem deutlich höheren Niveau als die Coachingaufgaben eingestuft wurden. Das heißt, in ihrem Falle gelangten die Experten zu dem Urteil, daß bei diesen Aufgaben weniger „Informationen zu Ausgangs- und Randbedingungen, zum Lösungsweg und zum angestrebten Ergebnis“ (Items 1, 2, 3) zur Verfügung standen. Demgegenüber entsprach das zur Lösung der verschiedenen Aufgabengruppen notwendige „spezifische Fachwissen“ (Dimension „Erforderliches Wissen“) der angestrebten Sequenzierung der Aufgabenanforderungen. Nach Urteil der Experten verlangten die drei Aufgabengruppen mit geringen Unterschieden nur „wenig bis mittleres Wissen über allgemeine Lösungsstrategien“ (Item 3). Hinsichtlich der Dimension „Freiheitsgrade“ entsprach die Beurteilung „zweckmäßiger Lösungen und Lösungswege“ (Item 1, 2) bei geringen bis mittleren Ausprägungen der im Programm vorgenommenen Sequenzierung. Ähnliches läßt sich auch für die Dimension „Widersprüchliche Ziele“ festhalten, insoweit die Anzahl der in den Aufgaben allgemein auftretenden „Ziele“ von den Coaching- über die Scaffolding- zu den Explorationsaufgaben hin zunahm. Auch bezüglich „wider-

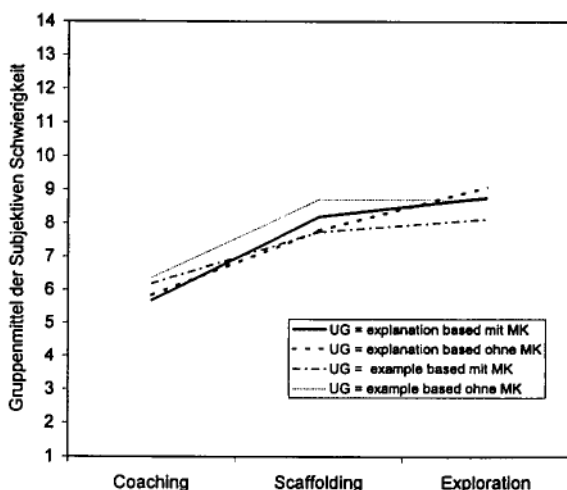
sprüchlicher Ziele“ ist eine vergleichbare Abstufung zu erkennen, allerdings mit wesentlich geringeren Ausprägungen.

Zusammenfassend kann festgestellt werden, daß die Problem- und Aufgabenanalyse von Hacker et al. (1998) hier als ein adäquates Verfahren beurteilt werden kann, um die Schwierigkeit von Lernaufgaben differenzierter zu betrachten, als es im Ansatz der Cognitive Apprenticeship gefordert wird. Bei hoher Übereinstimmung des Urteils konnte die Bedeutsamkeit der Dimension *Komplexität* als zentral für den Anforderungsgehalt volkswirtschaftlicher Lernaufgaben nachgewiesen werden. Die im Lehrprogramm enthaltenen Aufgaben sind (außer den obengenannten) Ausnahmen hinsichtlich ihrer Komplexität sequenziert.

### *Subjektive Aufgabenanalyse*

Zur subjektiven Einschätzung der Aufgabenschwierigkeit wurden in den Studien 1 und 2 zwei unterschiedliche Fragebögen eingesetzt. Der in der 1. Studie eingesetzte Fragebogen genügte nicht den Testgütekriterien und differenzierte nur auf einer Skala von 1 bis 5. Die mittlere subjektive Aufgabenschwierigkeit der Scaffoldingaufgaben betrug 2.42 ( $s = 0.52$ ), bei den Explorationsfragen betrug der Mittelwert 2.73 ( $s = 0.58$ ). Ein statistisch signifikanter Unterschied zwischen diesen Aufgabengruppen wurde knapp verfehlt. Die Berechnung von Korrelationen zwischen Schwierigkeitseinschätzungen und den Scaffolding- bzw. Explorationsleistungen ergab keine signifikanten Zusammenhänge. Aus Skepsis gegenüber der Validität dieses Fragebogens wurde in der 2. Studie die Zwei-Ebenen-Intensitätsskala (ZEIS) von Käßler & Pitrella (1988) eingesetzt. Mit der ZEIS-Skala liegt bei geringem methodischen Aufwand ein zuverlässiges Instrument hinsichtlich der Validität des zu untersuchenden Konstruktes der mentalen Beanspruchung sowie auch bezüglich weiterer Gütekriterien vor. Die Ergebnisse dieser subjektiven Beurteilung sind in der folgenden Abbildung dargestellt:

Abbildung 6: Subjektive Schwierigkeitseinschätzung in der 2. Studie



Bezüglich der fünf als „ziemlich leicht“ eingeschätzten Coachingaufgaben, der als „kaum leicht bis kaum schwierig“ eingeschätzten Scaffoldingaufgaben und als der als „etwas schwierig“ beurteilten Explorationsaufgaben erwiesen sich die zwischen den vier Untersuchungsteilgruppen feststellbaren Unterschiede als statistisch nicht signifikant. Insgesamt lag das Niveau der mentalen Beanspruchung im unteren Skalenbereich, Beurteilungen wie „schwierig“ traten kaum auf (vgl. Pfendler, 1993). Die Untersuchungsteilnehmer fühlten sich, nach diesen Daten zu urteilen, im allgemeinen mental nicht überfordert.

Die subjektiven Schwierigkeitseinschätzungen der drei Aufgabengruppen erwiesen sich zwischen Coaching- und Scaffolding- sowie zwischen Coaching- und Explorationsaufgaben als signifikant different (Wilcoxon-Test). Auch der Unterschied zwischen Scaffolding- und Explorationsaufgaben erwies sich nach diesem Test auf dem 5 %-Niveau als signifikant. Wir können daher schließen, daß die mentale Beanspruchung durch die Aufgabenstellungen für die Individuen aller Untersuchungsgruppen von Aufgabengruppe zu Aufgabengruppe zunahm. Es läßt sich zeigen, daß dieses Resultat nicht auf die unterschiedlichen Methoden, sondern hauptsächlich auf den steigenden kognitiven Anforderungsgehalt zurückführbar ist, der wesentlich durch die zunehmende Komplexität der Aufgaben bedingt wird (vgl. die Befunde der objektiven Aufgabenanalyse). Der steigende kognitive Anforderungsgehalt der Lernaufgaben (Merkmalsdimension *Sequenzierung*) korrespondiert mit der zunehmenden subjektiven mentalen Beanspruchung der Lernenden. Da die vier Untersuchungsgruppen trotz gleicher mentaler Beanspruchung auch in den objektiven Leistungsdaten keine statistisch signifikanten Unterschiede aufwiesen (s.u.), kann begründet davon ausgegangen werden, daß die verschiedenen realisierten Lehrstrategien weder objektive Leistungsunterschiede noch subjektive Beanspruchungsunterschiede bewirkt haben. In diesem Zusammenhang ist noch zu erwähnen, daß die objektiven Leistungsdaten in einem umgekehrt proportionalen Zusammenhang mit der mentalen Beanspruchung standen, da eine hohe mentale Beanspruchung mit schlechteren Leistungen und umgekehrt korrelierte: Als Korrelation der subjektiven Beanspruchung mit der Fehlerzahl in Coaching ergab sich der Wert  $r = .469$ , mit den erzielten Punktwerten in Scaffolding  $r = -.074$  und mit denen in Exploration  $r = -.379$ .

## 4.2 Ergebnisse der formativen Evaluation

In Anbetracht der zentralen Bedeutung des bereichsspezifischen Vorwissens wie auch der Lernmotivation von Schülern und ihrer Einstellung zu einem Schulfach unterzogen wir zunächst die

### *Eingangsbedingungen der Lernenden*

einer Analyse. Das *Vorwissen* der Untersuchungsteilnehmer wurde anhand der Leistungen in der „Einstiegsaufgabe“ in *Modeling*, die ökonomisches Hintergrundwissen zur Erzeugung eines unmittelbaren Verständnisses der

Geldpolitik der Bundesbank voraussetzt, und in der 2. Studie zusätzlich mit dem Wissenstest zu „Dimensionen der ökonomischen Bildung“ nach Beck (1993) erfaßt. Der Beck'sche Test zielt auf die Feststellung deklarativen Wissens in den Bereichen Mikro- und Makroökonomie. Die Ergebnisse dieser Messung sind in der folgenden Tabelle wiedergegeben:

Tabelle 3: Vorwissen und Homogenität der Untersuchungsgruppen

	1. Studie (UG = 2)		2. Studie (UG = 4)			
Untersuchungsgruppen	UG 1	UG 2	UG 1	UG 2	UG 3	UG 4
Einstiegsaufgabe MW [%]	31,11	29,76	14,97	14,97	22,08	12,86
Einstiegsaufgabe S	19,79	18,70	13,15	12,35	15,73	12,17
Wissenstest MW [Pkt.]	-	-	14,29	13,86	14,41	12,65
Wissenstest S	-	-	3,03	3,32	3,36	4,18

Anmerkung: 1. Studie: UG 1 = Erklärungsbegründet ohne Scaffolding; UG 2 = Erklärungs-begründet mit Scaffolding; 2. Studie: UG 1 = Erklärungs-begründet mit Metakognition; UG 2 = Erklärungs-begründet ohne Metakognition; UG 3 = Beispielbegründet mit Metakognition; UG 4 = Beispielbegründet ohne Metakognition.

Während die relative Leistung bei der „Einstiegsaufgabe“ in der 1. Studie (Studenten) bei einem Mittelwert von 30.46 ( $s = 18.93$ ) lag, fanden wir in der 2. Studie (Gymnasiasten) eine relative Leistung der Gesamtstichprobe mit Mittelwert 16.38 ( $s = 13.70$ ). Damit wird deutlich, daß das Ausgangsniveau der Leistungen von Studenten und Gymnasiasten bereits in der Einstiegsaufgabe signifikant unterschiedlich war, was bei der Interpretation von Vergleichen beider Studien zu berücksichtigen ist. Dagegen erwiesen sich die Unterschiede zwischen den einzelnen Untersuchungsgruppen als nicht signifikant (1. Studie: Mann-Whitney-U-Test  $p = .80$ ; 2. Studie: Kruskal-Wallis-Test  $p = .198$ ). Festzuhalten ist noch, daß in der 2. Studie die Gruppe „Beispielbegründet mit Metakognition“ das beste Ausgangsleistungsniveau (Mittelwert 22.08;  $s = 15.73$ ) erzielte. Beim Vergleich der beiden „Extremgruppen“ („Beispielbegründet mit und ohne Metakognition“) ist ein mäßig signifikanter Unterschied ( $p = .052$  nach Kruskal-Wallis-Test) feststellbar.

Die Items des eingesetzten Tests sind Rasch-skaliert, so daß aus inhaltlichen und versuchsorganisatorischen Gründen auf die Wissensteilbereiche Mikro- und Makroökonomie zugegriffen werden konnte. Im Vergleich zur Normstichprobe ( $N = 670$  Gymnasiasten allgemeinbildender Gymnasien aus mehreren Bundesländern) können alle in den Untersuchungsgruppen erreichten Leistungen als durchschnittlich eingestuft werden. Der Mittelwert der Stichprobe der 2. Studie lag bei 13.83 Punkten ( $s = 3.5$ ) von 26 maximal erreichbaren Punkten. Die zwischen den Untersuchungsgruppen auftretenden Leistungsunterschiede erwiesen sich allesamt als statistisch nicht signifikant (Kruskal-Wallis-Test:  $p = .462$ ). In der 2. Studie wurde auch der Einfluß der



verschiedenen Gymnasien, aus denen die Untersuchungsteilnehmer stammten, als mögliche Kovariable untersucht. Weder für die „Einstiegsaufgabe“ (Kruskal-Wallis-Test:  $p = .149$ ) und den Beck’schen Wissenstest (Kruskal-Wallis-Test:  $p = .333$ ) noch für die erfaßten Einstellungen zu Wirtschaftsthemen (EWF) nach Beck (1993) (Kruskal-Wallis-Test:  $p = .553$ ) konnten überzufällige Unterschiede festgestellt werden.

Zusammenfassend ist festzuhalten, daß zwischen Studenten und Gymnasiasten hinsichtlich des fachspezifischen Vorwissens ein bedeutsamer Niveauunterschied feststellbar war. In beiden Evaluationsstudien traten im Vorwissen allerdings keine bedeutsamen Unterschiede zwischen den verschiedenen Untersuchungsgruppen auf.

Hinsichtlich der Voraussetzungen in motivationaler Hinsicht erbrachte die Eingangsmessung folgende Daten:

Tabelle 4: Motivation und Einstellung der Untersuchungspartner

	1. Studie (UG = 2)		2. Studie (UG = 4)			
Untersuchungsgruppen	UG 1	UG 2	UG 1	UG 2	UG 3	UG 4
LMT MW [T-Werte]	47,91	53,26	-	-	-	-
LMT S	9,26	10,70	-	-	-	-
EWF MW [Pkt.]	-	-	3,25	3,42	3,36	3,23
EWF S	-	-	0,50	0,46	0,51	0,63

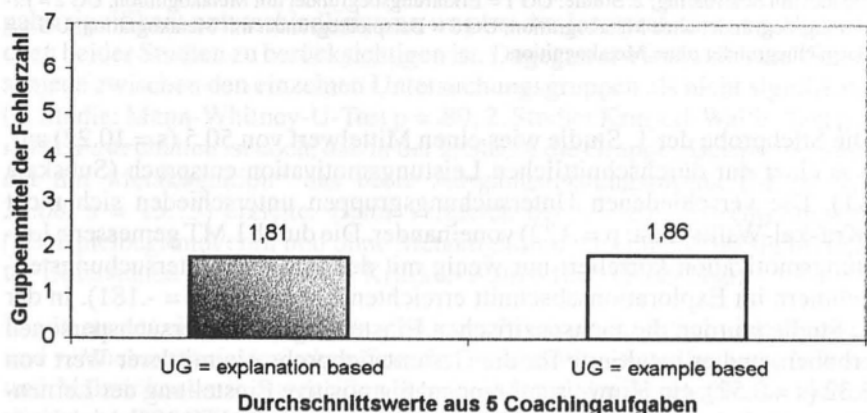
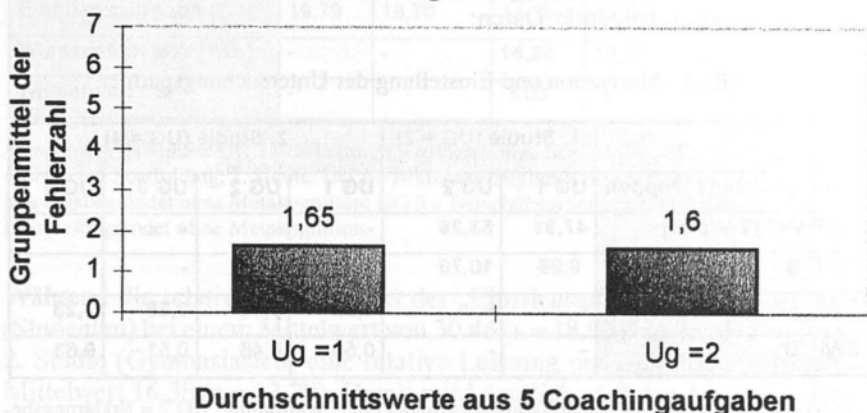
Anmerkung: 1. Studie: UG 1 = Erklärungsbegründet ohne Scaffolding; UG 2 = Erklärungs-begründet mit Scaffolding; 2. Studie: UG 1 = Erklärungs-begründet mit Metakognition; UG 2 = Erklärungs-begründet ohne Metakognition; UG 3 = Beispielbegründet mit Metakognition; UG 4 = Beispielbegründet ohne Metakognition.

Die Stichprobe der 1. Studie wies einen Mittelwert von 50.5 ( $s = 10.22$ ) auf, was einer nur durchschnittlichen Leistungsmotivation entsprach (Subskala L1). Die verschiedenen Untersuchungsgruppen unterschieden sich nicht (Kruskal-Wallis-Test:  $p = .122$ ) voneinander. Die durch LMT gemessene Leistungsmotivation korreliert nur wenig mit den von den Untersuchungsteilnehmern im Explorationsabschnitt erreichten Leistungen ( $r = -.181$ ). In der 2. Studie wurden die fachspezifischen Einstellungen der Versuchspersonen erhoben, und es resultierte für die Gesamtstichprobe ein mittlerer Wert von 3.32 ( $s = 0.52$ ), ein Hinweis auf eine mäßig positive Einstellung der Lernenden bezüglich ökonomischer Themen, da die von Beck (1993) berichtete Norm für Allgemeinbildende Gymnasien bei einem Mittelwert von 2.67 liegt. Signifikante Unterschiede zwischen den Untersuchungsgruppen waren auch in bezug auf die Einstellungen nicht nachweisbar (Kruskal-Wallis-Test:  $p = .574$ ). Bemerkenswerterweise betrug die Korrelation der Einstellungen zu wirtschaftlichen Fragen mit dem fachspezifischen Wissen  $r = .293$  gegenüber einer Korrelation von  $r = .46$  in der Normstichprobe Becks.

## Lernleistungen in Coaching

Nach Bearbeitung des Modeling hatten die Versuchspersonen beim Coaching zum ersten Mal konstruktiv Lernaufgaben zu bearbeiten, bei denen es um die Übertragung des Regelkreismodells auf wirtschaftliche Fragestellungen ging. Die Aufgabenstellung bestand im Zuordnen ausgewählter Begriffe in ein vorbereitetes Regelkreisschema. Es handelte sich also um eine übersichtliche, gut strukturierte Aufgabe, bei der nach drei aufeinanderfolgenden falschen Zuordnungen dem Lernenden eine individuelle Hilfestellung und Korrektur gegeben wurde. Pro Aufgabe waren jeweils sieben Begriffe zuzuordnen, was der maximal möglichen Fehlerzahl entspricht.

Abbildung 7:  
Mittlere Fehler in Coaching der 1. Studie und 2. Studie



Anmerkung: UG Beispielsbegründet: N1 = 41; UG Erklärungs begründet: N2 = 41

In Studie 1 bearbeiteten beide Untersuchungsgruppen bis zum Ende des Coachingabschnittes ein vollkommen identisches Lernprogramm, deshalb wurden keine Unterschiede erwartet. Insgesamt kann festgestellt werden, daß

die Versuchspersonen alle Coaching-Aufgaben mit überdurchschnittlichem Erfolg bearbeiteten und kaum Schwierigkeiten bei der Lösung qua Analogiebildung hatten. D. h. es gelang ihnen, das Regelkreismodell erfolgreich auf neue volkswirtschaftliche Sachverhalte zu übertragen. Die subjektive Einschätzung des Neuigkeitsgehaltes der verwendeten Konzepte ergab, daß das Regelkreismodell (nach den Instruktionen in *Modeling*) den größten Bekanntheitsgrad aufwies, gefolgt vom Geldkreislauf, den Aufgaben der Bundesbank und der Geldmengenbestimmung bis hin zur Buchgeldschöpfung, die den Untersuchungsteilnehmern am wenigsten vertraut war. Bei keiner abhängigen Variable (Fehlerzahl, Bearbeitungszeit und Anzahl der Lernschritte) resultierten in Studie 1 statistisch signifikante Unterschiede zwischen den Untersuchungsgruppen.

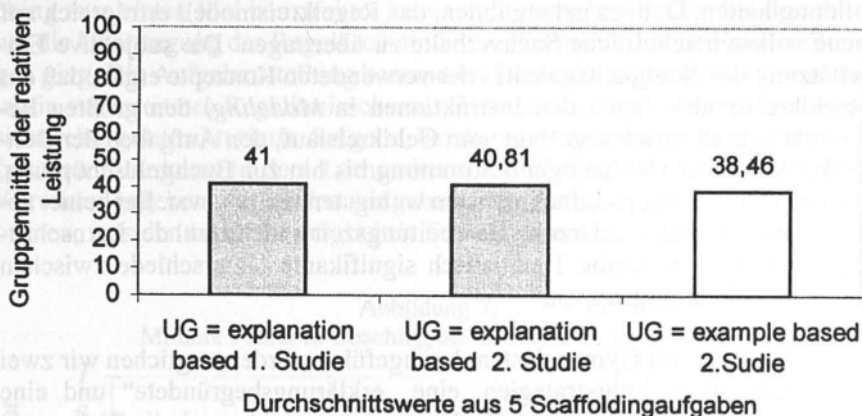
In Studie 2, die mit Gymnasiasten durchgeführt wurde, verglichen wir zwei verschiedenartige Lehrstrategien, eine „erklärungsbegründete“ und eine „beispielbegründete“, miteinander. Insgesamt erreichte auch die Zielgruppe der Gymnasiasten ein überdurchschnittliches Leistungsniveau beim *Coaching*. Die mittlere Fehlerzahl lag bei gleicher Schrittzahl und einer etwa 30 Sekunden längeren Bearbeitungszeit gerade zwei Zehntel höher als bei den Studenten, wobei auch keine Unterschiede zwischen der „erklärungsbegründeten“ und der „beispielbegründeten“ Lehrstrategie feststellbar waren. Somit erwiesen sich die Lernleistungen von Studenten und Gymnasiasten in den Coachingaufgaben als annähernd wertgleich, wobei noch darauf hinzuweisen ist, daß beide Lerngruppen überdurchschnittliche Leistungen erbrachten. Auch das Lernziel, Analogieschlüsse in eindeutig definierten Problemstellungen korrekt zu ziehen, wurde in beiden Studien vollständig erreicht. Darüber hinaus erwiesen sich die beiden zugrundegelegten Lehrstrategien als durchaus vergleichbar in ihrer Effektivität. Dies konnte auch schon in der Voruntersuchung festgestellt werden (vgl. Seel et al., 1998).

### *Lernleistungen beim Scaffolding*

Das zweite innerhalb des Lehrprogramms „Wirtschaftssysteme“ zu lösende Aufgabenpaket besteht aus *Scaffolding*-Aufgaben. Hierbei handelt es sich um teilweise strukturierte Problemstellungen mittlerer Komplexität (siehe Aufgabenanalyse), die als offene Fragen gestellt werden und dementsprechend schriftlich zu beantworten sind. Dabei können allgemeine Hilfestellungen wie Lexikon, Darstellungen des Geldkreislaufs, das Regelkreisschema oder auch Regelkreisbeispiele genutzt werden.

Die schriftlich gegebenen Antworten wurden nach der Anzahl richtig beschriebener Ursache-Wirkungszusammenhänge bewertet und an der maximal erreichbaren Gesamtpunktzahl relativiert. Der Auswertung wurde ein Beurteilungsschema zugrunde gelegt, das in Zusammenarbeit mit einem Experten, einem promovierten Volkswirt, erarbeitet worden war und das vor dem erneuten Einsatz in der 2. Studie überarbeitet wurde. In der folgenden Abbildung sind die entsprechenden Werte für die verschiedenen Versuchsbedingungen dargestellt.

Abbildung 8:  
Vergleich der Leistungen in 5 Scaffoldingfragen



Die weitergehende Auswertung der Lernleistungen erbrachte folgende Werte:

Tabelle 5:  
Auswertung der 5 Scaffoldingfragen

	1. Studie (UG = 2)		2. Studie (UG = 4)	
Untersuchungsgruppen	UG 1	UG 2	UG 1+UG 2	UG 3 +UG 4
Leistung; MW [%]	-	41,00	40,81	38,46
S	-	15,00	13,76	13,19
Hilfsmittelnutzung; MW	4 RK *	1 RK *	4,46	4,75
S	-	-	3,13	2,94

Anmerkung: 1. Studie: UG 1 = Erklärungsbe gründet ohne Scaffolding; UG 2 = Erklärungsbe gründet mit Scaffolding; 2. Studie: UG 1 = Erklärungsbe gründet mit Metakognition; UG 2 = Erklärungsbe gründet ohne Metakognition; UG 3 = Beispielbe gründet mit Metakognition; UG 4 = Beispielbe gründet ohne Metakognition.

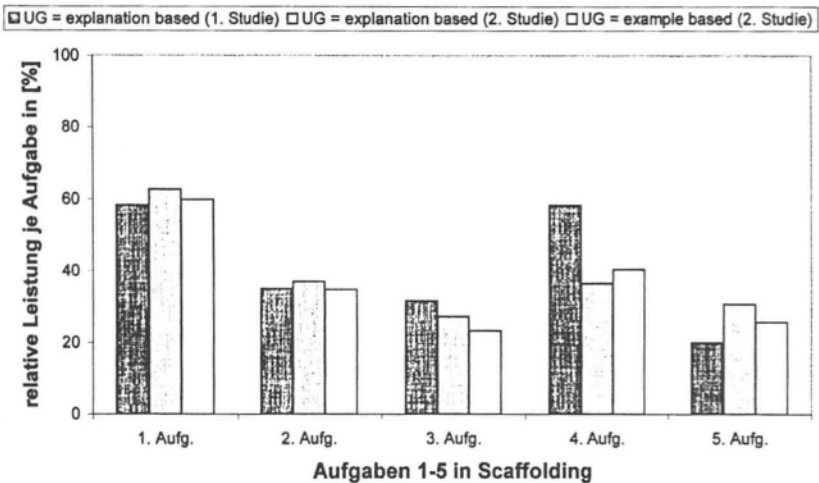
\* Anwendung der Regelkreisschemas (RK) bei Beantwortung der letzten Scaffoldingfrage.

Die Leistungen in der 1. Studie sind als gerade noch durchschnittlich zu bewerten. Angesichts der hohen Streuungen ist festzustellen, daß es auch Versuchspersonen gab, die überhaupt keinen Punkt erreichten. Die höchsten Leistungen traten bei den Aufgaben 1 und 4 mit 58.33 % und die geringsten Leistungen bei Aufgabe 5 mit 20.0 % (UG = 2) auf. Hinsichtlich der Nutzung der Hilfsmittel war festzustellen, daß lediglich fünf Lernende (von insgesamt 34) die Regelkreisdarstellung bei der Bearbeitung der letzten Lernaufgabe verwendeten. Dabei erwies sich der Zusammenhang mit der erreichten Leistung als statistisch signifikant ( $r = .4803$ ;  $p = .005$ ). Der Unterschied zwischen den Behandlungsgruppen bei dieser Lernaufgabe war ohne Signifikanz (UG 1 = 30.59 %; UG 2 = 20.0 %) – ein deutlicher Hin-

weis auf die geringe Lernwirksamkeit der beim Scaffolding angebotenen Hilfestellungen und Hilfsmittel.

In der 2. Studie wurden die Untersuchungsgruppen zu zwei Gruppen zusammengefaßt, da eine Gegenüberstellung der Untergruppe „mit versus ohne Metakognition“ erst nach Durchführung des Metakognitionsabschnittes, also im Explorationsabschnitt, als sinnvoll beurteilt wurde. Verglichen mit den Ergebnissen in Studie 1 war festzustellen, daß die Gymnasiasten in Studie 2 etwa das gleiche (mäßige) Leistungsniveau wie die Studenten erreichten:

Abbildung 9:  
Relative Leistung je Lernaufgabe beim Scaffolding



Auf die einzelnen Aufgaben bezogen, lag das Leistungsspektrum der Versuchspersonen in Studie 2 zwischen dem Höchstwert von 62.72 % bei der ersten Lernaufgabe in der Behandlungsgruppe „erklärungsbe gründet“ und dem Minimalwert von 23.33 % in der Behandlungsgruppe „beispielbe gründet“ bei der dritten Lernaufgabe. Bei Aufgabe 5 schnitten die Studenten der Untersuchungsgruppe UG = 2 (in Studie 1) mit 20.00 % sogar schlechter ab als die Gymnasiasten. Eine Gegenüberstellung der „erklärungs- und beispielbe gründeten Strategie“ ergibt bei nur geringem Vorsprung der erstgenannten Strategie weder in der Gesamtleistung noch in Einzelleistungen in den fünf Lernaufgaben statistisch bedeutsame Unterschiede.

Eine naheliegende Schlußfolgerung aus den dargestellten Ergebnissen besteht in der Annahme, daß Gymnasiasten unter vergleichbaren Lernbedin gungen das gleiche Leistungsniveau wie Studenten erreichen können. Be zieht man dabei das Vorwissen als Kovariate ein, kann für die Versuchsgrup pe der Gymnasiasten sogar ein größerer Lernzuwachs festgestellt werden. An dererseits deuten die Befunde darauf hin, daß beide Lerngruppen an Lei-

stutzungsgrenzen stießen, die zum Teil auf eine noch nicht optimale Programmgestaltung, insbesondere was die Umsetzung interaktiver Leistungsrückmeldungen und prozeßorientierter, das Problemlösen unterstützender Hilfestellungen, betrifft oder weitergehend auch darauf schließen läßt, daß Scaffolding womöglich nur eingeschränkt in eine multimediale Lernumgebung zu inkorporieren ist. Dieser Problematik gehen wir in unserer zur Zeit laufenden Untersuchung gezielt auf den Grund.

### Effektivität der metakognitiven Komponenten

Die Konzeption der Cognitive Apprenticeship sieht vor, daß nach *Scaffolding* ein kritisches Nachdenken der Lernenden darüber stattfinden soll, wie sie zuvor beim Problemlösen vorgegangen sind. Dies sollen sie in Form der Selbstverbalisierung (= *Artikulation*) und der *Reflexion* der eigenen Vorgehensweisen im Vergleich mit denen des Experten (beim *Modeling*) (wie auch anderer Lernender) realisieren. Wir fassen im folgenden beide Strategien metakognitiver Kontrolle zusammen. Erfafßt wurden sie (nur) in der 2. Studie auf der Grundlage entsprechender Fragen im Lehrprogramm. Die Antworten wurden programmseitig wortwörtlich protokolliert und dann unter Verwendung eines von Hasselhorn (1992 entwickelten Systems der Kategorisierung für Metakognitionen, d. h. hinsichtlich der Anzahl aufgeschriebener, unterschiedlicher Gedanken zur Metakognition) ausgewertet. Die Ergebnisse dieser Analyse sind für die beiden relevanten Untersuchungsgruppen in der folgenden Tabelle wiedergegeben:

Tabelle 6:  
Metakognitionskomponente in der 2. Studie

	2. Studie		
Untersuchungsgruppen	UG 1	UG 3	U-Test [ p ]
Metakog. Articulation MW	8,33	7,36	.705
S	5,80	3,49	
Metakogn. Reflection MW	2,67	2,27	.575
S	1,83	1,24	

Anmerkung: 2. Studie: UG 1 = Erklärungs begründet mit Metakognition; UG 3 = Beispielbegründet mit Metakognition

Man hätte erwarten können, das die unterschiedlichen Lehrstrategien auch ein mehr oder weniger intensives Metakognitionsverhalten auslösen, was aber nicht der Fall war. Andererseits kann dieses Ergebnis auch nicht erstauen, da zwischen den Lehrstrategien bisher keine Unterschiede auftraten. Bemerkenswerterweise konnten wir einen signifikanten Unterschied zwischen der Häufigkeit von Aussagen in *Artikulation* und derjenigen in *Reflexion* (Wilcoxon-Test:  $p = .001$ ) feststellen, insoweit sich die Versuchspersonen mehrheitlich häufiger zu Artikulation als zu Reflexion äußerten.

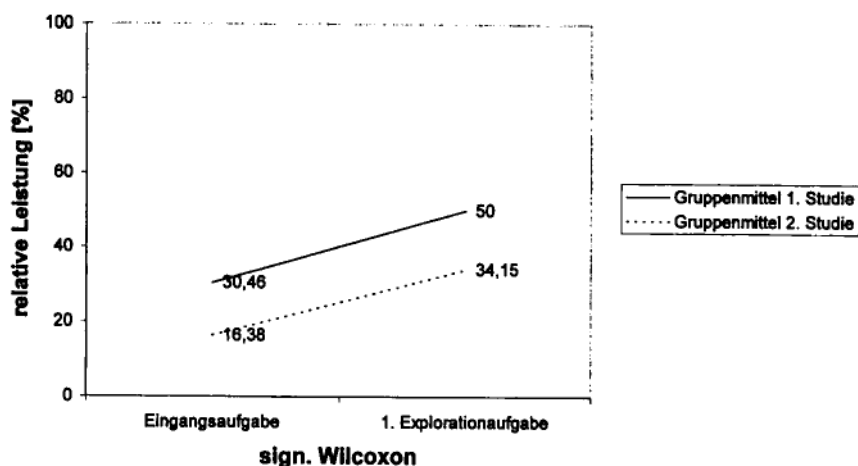
Die uns interessierende Frage nach Zusammenhängen der metakognitiven Komponenten mit den abschließenden Leistungen im Explorationsabschnitt kann vorerst nur sehr eingeschränkt beantwortet werden, da nur leichte Tendenzen festzustellen waren. Bildet man die Summe aus den Werten in Artikulation und Reflexion, ergibt sich eine bivariate Korrelation nach Pearson von  $r = .234$  ( $p = .127$ ). Hierbei gilt aber zu beachten, daß die Metakognitionskomponenten in dem Lehrprogramm „Wirtschaftssysteme“ zum Zeitpunkt von Studie 2 noch nicht optimal umgesetzt waren. Daher erwarten wir für die gerade laufende Evaluationsstudie andere Ergebnisse. Da die unzureichende Interaktivität des vorliegenden Multimedia-Programms ein effektives Reflektieren über die eigene Aufgabenbearbeitung nur in einem engen Rahmen erlaubt, werden die Komponenten der Metakognition, *Artikulation* und *Reflexion* in Form eines Gedankenaustauschs in kooperativen Lerngruppen (im sog. „teach-back“-Verfahren) geprüft.

### *Lernleistungen im Programmteil Exploration*

Die letzte innerhalb des Lehrprogramms zu lösende Aufgabengruppe bezieht sich auf die Exploration. Sie stellt die Anforderung, daß die Lernenden weitgehend selbständig und umfassend neue Aufgaben höherer Komplexität bearbeiten und lösen. In dem Lehrprogramm „Wirtschaftssysteme“ wird den Lernenden erlaubt, eine Vielzahl an Hilfestellungen (z. B. Lexikon, Geldkreislauf oder Regelkreisbeispiele) zu benutzen.

Wie bei den Scaffolding-Aufgaben wurden die schriftlich oder unter Verwendung eines Regelkreisschemas gegebenen Antworten nach der Anzahl korrekt beschriebener Ursache-Wirkungszusammenhänge bewertet und gegen die maximal erreichbare Gesamtpunktzahl relativiert. Es zeigte sich, daß die Leistungen in allen Behandlungsgruppen unter dem Niveau der bei den Scaffoldingaufgaben erzielten Leistungen lagen.

Abbildung 10:  
Vergleich von Einstiegs- und 1. Explorationsaufgabe





Die erste Explorationsaufgabe ist inhaltlich identisch mit der zu Beginn des Lernprogramms gestellten Lernaufgabe („Einstiegsaufgabe“), so daß ein Vergleich beider Aufgaben am ehesten Auskunft über die Lerneffektivität des Programms gibt. Dieser Vergleich anhand der relativen Leistungen bestätigt einen deutlichen Leistungsanstieg in beiden Studien, der auf die Bearbeitung des Lernprogramms zurückgeführt werden kann. Die Prüfung mittels Wilcoxon-Test für abhängige Stichproben ergab einen signifikanten Unterschied in beiden Studien (1. Studie:  $p .005$ ; 2. Studie:  $p .001$ ). In bezug auf die übrigen Aufgaben des Programmteils Exploration resultierten folgende Kennwerte:

Tabelle 7:  
Leistungen in den 5 Explorationsfragen

	2. Studie		
Untersuchungsgruppen	UG 1	UG 3	U-Test [ p ]
Metakog. Articulation MW	8,33	7,36	.705
S	5,80	3,49	
Metakogn. Reflection MW	2,67	2,27	.575
S	1,83	1,24	

Anmerkung: 1. Studie: UG 1 = Erklärungsbegründet ohne Scaffolding; UG 2 = Erklärungsbe-  
gründet mit Scaffolding; 2. Studie: UG 1 = Erklärungsbegründet mit Metakognition; UG 2 = Er-  
klärungsbe- gründet ohne Metakognition; UG 3 = Beispielbegründet mit Metakognition; UG 4 =  
Beispielbegründet ohne Metakognition.

In der 1. Studie traten zwischen den Untersuchungsgruppen keine statistisch bedeutsamen Unterschiede bei den Explorationsleistungen insgesamt, wie auch bei den einzeln betrachteten fünf Fragen auf. Lediglich bei der zweiten Explorationsfrage trat ein Unterschied auf dem 10 %-Niveau zugunsten der Untersuchungsgruppe 1 auf. Diese Ergebnisse lassen den Schluß zu, daß die Variation der Versuchsbedingungen in Scaffolding keinen Einfluß auf die Leistungen im Explorationsabschnitt hatte.

Die Versuchspersonen der 1. Studie erreichten etwa das gleiche Leistungsniveau bei den Scaffolding- und den Explorationsaufgaben. Demgegenüber erzielten die Gymnasiasten (in Studie 2) einen größeren Lernzuwachs, wenn man ihre geringeren Vorkenntnisse berücksichtigt. Tatsächlich erreichte das Leistungsniveau der Gesamtstichprobe der Gymnasiasten einen Mittelwert von 32.55 ( $s = 11.88$ ), was dem Leistungsniveau der Studenten entsprach und trotzdem nur als durchschnittlich zu beurteilen war. Möglicherweise ist dies auf die unzureichende Nutzung der angebotenen Hilfsmittel zurückzuführen, insofern die Versuchspersonen im Explorationsabschnitt noch seltener als beim Scaffolding von diesen Gebrauch machten. Gestützt wird diese Annahme durch die Beobachtung, daß die zielorientierte, kognitive Tätig-



keit der Nutzung des Notizblocks in Form selbst eingetragener Stichpunkte nur knapp die Signifikanzgrenze ( $r = .265$ ;  $p = .149$ ) verfehlte.

### 4.3 *Ergebnisse der summativen Evaluation*

Die Gesamteffektivität des multimedialen Lehrprogramms „Wirtschaftssysteme“ i. S. einer summativen Evaluation wurde zunächst auf der Grundlage einer Parallelform des Wissenstest nach Beck (1993; Form B) gemessen, um den Wissensstand zum entsprechenden Zeitpunkt zu ermitteln. Es zeigte sich, daß zwischen Vor- und Nachtest ein mittlerer Zuwachs von 1.4 Punkten erreicht wurde. Ein Lerneffekt der Arbeit mit dem Lehrprogramm konnte damit für das deklarative Wissen nicht nachgewiesen werden. Auch traten keine statistisch bedeutsamen Unterschiede zwischen den einzelnen Untersuchungsgruppen auf. Auf den ersten Blick scheint dieses Ergebnis gegen eine Lerneffektivität des Multimedia-Programms „Wirtschaftssysteme“ zu sprechen, doch ist dies zu relativieren, da der Beck'sche Wissenstest nicht kriterienorientiert ist und daher nicht unmittelbar mit den Zielsetzungen des Lehrprogramms korrespondiert. Daher spricht das Ergebnis unseres Nachtests für die Persistenz und Reliabilität des Beck'schen Tests. Zur Messung programmspezifischer Lerneffekte scheint er demgegenüber ungeeignet zu sein. Dieses Urteil wird auch durch die Ergebnisse unserer Nachuntersuchung gestützt, die drei Monate nach Beendigung der Lernphase zur Ermittlung der zeitlichen Stabilität des Wissens durchgeführt wurde. Hierbei wurde ein aus der Kombination von Items mittlerer Schwierigkeit der Formen A und B des Wissenstests von Beck entwickelter Test (Form C genannt) verwendet. Wieder lag der Gesamtmittelwert nur etwa 1.3 Punkte höher, was auf stabile Leistungen hinwies. Berücksichtigt man nun noch, daß das Wissensgebiet in der Zwischenzeit auch im Unterricht behandelt worden war, so kommt man nicht umhin, dem Beck'schen Wissenstest eine erstaunliche Stabilität zuzuschreiben, für die Messung kurz- oder langfristiger Lerneffekte ist er demgegenüber jedoch nicht geeignet.

In Anbetracht dessen führten wir zum Zwecke einer summativen Evaluation der Effektivität des Lehrprogramms Regressionsanalysen durch, um ein entsprechendes theoretisches Modell multipler Beeinflussungsmöglichkeiten der Leistungen im Explorationsabschnitt zu prüfen. Als unabhängige Variablen wurden zugrundegelegt: die Persönlichkeitsvariablen LPS UT10, SPM (nur 1. Studie), KFT 3+4, LMT (2. Studie: EWS) und FKK 1-4 sowie die erreichten Leistungen in der Einstiegsaufgabe, den Coachingaufgaben und in der letzten Scaffoldingaufgabe. In Studie 2 gaben wir die Werte der fünf Scaffoldingfragen wie auch die des Wissenstests von Beck (1993) und der festgestellten EPI-Werte ein. Die zu prüfende Hypothese lautete, daß eine wirksame Lehrmethode die Zielgröße „Explorationsleistung“ vorrangig durch zuvor erfolgreiche Lernleistungen und nicht durch Persönlichkeitsvariablen erklären müßte. Diese Annahme konnte durch die Ergebnisse einer multiplen Regressionsanalyse weitgehend bestätigt werden.

Tabelle 8:  
Ergebnisse der multiplen Regressionsanalyse, Methode schrittweise

Abhängige Variable	Welche Studie	Unabhängige Variable	Stand. Regressionskoeffizient [Beta]	Korr. Bestimmtheitsmaß [R <sub>korr</sub> ]
Exploration gesamt	1. Studie	Scaffolding Frage 8 LPS	.601 .315	.387
	2. Studie	mittl. Fehler im Coaching Scaffolding Frage 2	-.450 .312	.291
Frage 1	1. Studie	Scaffolding Frage 8 FKK	.526 .446	.363
	2. Studie	Vorwissenstest Scaffolding Frage 2	.446 .411	.371
Frage 2	1. Studie	Scaffolding Frage 8 LPS	.546 .438	.411
	2. Studie	mittl. Fehler im Coaching	-.420	.155
Frage 3	1. Studie	Scaffolding Frage 8	.563	.291
Frage 4	1. Studie	SPM Scaffolding Frage 8	.467 .399	.350
	1. Studie	mittl. Fehler im Coaching	-.460	.180
Frage 5	2. Studie	Scaffolding Frage 5 Einstiegsaufgabe	.495 -.307	.199
	1. Studie	-	-	-
Frage 6	2. Studie	EPI/ Extraversion	-.311	.073
	1. Studie	LPS	.462	.182
Frage 7	2. Studie	mittl. Fehler im Coaching	-.452	.183

Die Leistungen in den Lernaufgaben der Komponente *Exploration* lassen sich größtenteils durch erfolgreich absolvierte Etappen der Cognitive Apprenticeship wie des Coaching oder Scaffolding vorhersagen. In der 1. Studie treten mehrmalig Effekte der Intelligenztests LPS und SPM auf, was darauf hinweist, daß auch die Intelligenz einen wenn auch vergleichsweise geringen Einfluß auf die Leistungen in den Explorationsaufgaben ausübt. Dieser Einfluß war bei den Gymnasiasten in der 2. Studie nicht mehr erkennbar. Insgesamt können die Ergebnisse der multiplen Regressionsanalyse als eine tendenzielle Stützung der Annahmen zur Effektivität der Cognitive Apprenticeship gewertet werden. Der relativ niedrige Betrag des korrigierten Bestimmtheitsmaßes zeigt jedoch, daß es noch andere relevante Einflußgrößen auf die Leistungen in Exploration geben muß, die über die bisher einbezogenen Variablen hinausgehen.

## 5. Untersuchungskritik und Konsequenzen

Obwohl die vorliegenden Befunde grundsätzlich auf eine Lerneffektivität des nach den Prinzipien der Cognitive Apprenticeship gestalteten Multimedia-Programms „Wirtschaftssysteme“ hindeuten, sind abschließend einige kritische Punkte anzumerken:

1. Unsere Evaluationsdaten lassen die Vermutung zu, daß die Versuchspersonen beider Untersuchungsdurchgänge durchaus noch über Leistungsreserven verfügten. Diese konnten sie jedoch nicht voll ausschöpfen, da die Gestaltung des Lernprogramms im Blick auf das Konzept der Cognitive Apprenticeship noch deutliche Schwächen aufwies. Zwar waren bei Durchführung der beiden in diesem Beitrag geschilderten Evaluationsstudien die einzelnen Methoden der Cognitive Apprenticeship (*modeling, coaching, scaffolding, metacognition* und *exploration*) bereits in das Lehrprogramm inkorporiert, aber ihre Umsetzung auf der inhaltlichen und methodischen Ebene ließ, wie die objektive Aufgabenanalyse zeigt, noch zu wünschen übrig. Daher unterzogen wir das gesamte Multimedia-Programm auf der Grundlage der vorliegenden Evaluationsdaten einer grundlegenden Modifikation und Revision, indem wir z. B. eine Vermittlung von Grundlagen- und Begriffskenntnissen im Modeling-Abschnitt über mehrere Sinneskanäle mittels anschaulicher Modelle, Audio- und Videosequenzen zur besseren Veranschaulichung des Lehrstoffes anzielten ebenso, wie wir uns eine Verbesserung der Qualität der Metakognitionskomponenten durch Herausheben dieses Abschnittes aus der multimedialen Lernumgebung und stattdessen eine Realisierung in Form strukturierter Gruppenarbeit versprechen (vgl. dazu auch die entsprechenden Schlußfolgerungen von Casey, 1996).
2. Die Versuchspersonen arbeiteten an zwei aufeinanderfolgenden Abenden, also nach absolviertem Schultag und häufig noch zusätzlichen Lehrangeboten am Nachmittag und leisteten für unsere Untersuchung quasi Überstunden. Hinzu kam noch, daß die Arbeitszeit, je nach individuellem Arbeitstempo bis zu 6 Stunden pro Tag betragen konnte. Es ist anzunehmen, daß sich diese Bedingungen auf die Konzentrations- und Leistungsfähigkeit der Untersuchungspartner auswirkten, auch wenn dies nicht aus den Nachbefragungen deutlich wurde.
3. Es wurden Gruppenmittelwerte je Untersuchungsgruppe gebildet, bei denen die meisten Kriterien recht hohe Standardabweichungen aufwiesen. Die inter- wie auch intraindividuellen Leistungsunterschiede fanden bislang zu wenig Berücksichtigung, ihnen ist in der nächsten Evaluationsstudie gezielt auf den Grund zu gehen.
4. Die Dimension „sozio-kultureller Kontext“ der Cognitive Apprenticeship zielt durch den authentischen Kontext auf eine hohe Lernmotivation. Der Lehrinhalt konnte bisher nur über schlußfolgerndes Denken auf persönliche Handlungsfelder übertragen werden. Die überarbeitete Version des Lehrprogramms nimmt stärker auf diese Dimension Bezug, die in der EWU bevorstehende Einführung des EURO wird als motivationsfördernder Handlungsrahmen und auch für das Transferproblem im Programmabschnitt „Exploration“ genutzt. Dadurch dürften jedem Lernenden persönliche Berührungspunkte zum Lehrinhalt einsichtig werden.

## Literatur

- Ahn, W.K., Brewer, W.F., & Mooney, R.J. (1992). Schema acquisition from a single example. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18 (2), 391-412.
- Al-Diban, S., Held, S., Hess, C., & Seel, N.M. (i. pr.). Wie Schüler einen abstrakten Gegenstandsbereich zu meistern lernen. Didaktisches Design des Multimediaprogramms Dynamische Wirtschaftssysteme. *Zeitschrift für das berufliche Bildungswesen*.
- Beck, K. (1993). *Dimensionen der ökonomischen Bildung. Meßinstrumente und Befunde*. Nürnberg: Universität (Abschlußbericht zum DFG-Projekt 'Wirtschaftskundlicher Bildungs-Test [WBT]. Normierung und internationaler Vergleich').
- Bransford, J.D., Vyl, N., Kinzer, C., & Risko, V. (1990). Teaching, thinking and content knowledge: Toward an integrated approach. In B.F. Jones & L. Idol (Bds.), *Dimensions of thinking and cognitive instruction: Implications for educational reform*.
- Casey, C. (1996). Incorporating cognitive apprenticeship in multi-media. *Educational Technology Research and Development*, 44 (1), 71-84.
- Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology*, 87 (1), 5-17.
- Choi, J.I., & Hannafin, M. (1995). Situated cognition and learning environments: Roles, structures, and implications for design. *Educational Technology Research and Development*, 43 (2), 53-69.
- Cognition and Technology Group at Vanderbilt (1997). *The Jasper project. Lessons in curriculum, instruction, assessment, and professional development*. Hillsdale, NJ: Lawrence Erlbaum.
- Collins, A.S., Brown, J.S., & Holum, A. (1993). Cognitive apprenticeship: Making thinking visible. *American Educator*, 15 (3), 6-11, 38-46.
- Collins, A., Brown, J.S. & Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L.B. Resnick (Ed.) (1989), *Knowing, learning, and instruction* (pp. 453-494). Hillsdale, NJ: Erlbaum.
- Dijkstra, S., Seel, N.M., Schott, F. & Tennyson, R.D. (1997), *Instructional design: International perspectives. Vol. 2: Solving instructional design problems*. Mahwah, NJ.: Erlbaum.
- Dugdale, S. (1994). Using students' mathematical inventiveness as a foundation for software design: Toward a tempered constructivism. *Educational Technology Research and Development*, 42 (1), 57-73.
- Farmer, J.A., Buckmaster, A., & Legrand, A. (1992). Cognitive apprenticeship: Implications for continuing professional education. *New Directions for Adult and Continuing Education*, 55, 41-49.
- Flechsig, K.H. (1990). *Einführung in CEDID. Ein tätigkeitsunterstützendes und wissensbasiertes System für computerergänzt didaktisches Design*. Göttingen: CEDID GmbH.
- Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39, 93-104.
- Goodyear, P. (1997). Instructional design environments: Methods and tools for the design of complex instructional systems. In S. Dijkstra, N.M. Seel, F. Schott & R.D. Tennyson (Eds.), *Instructional design: International perspectives. Vol. 2: Solving instructional design problems* (pp. 83-111). Mahwah, NJ.: Erlbaum.
- Hacker, W., Sachse, P. & Schroda, F. (1998). Design thinking – possible ways to successful solutions in product development. In H. Birkhofer, P. Badke-Schaub & E.

- Frankenberger (Eds.), *Designers – the Key to successful Product Development*. London: Springer.
- Hannafin, M.J. (1992). Emerging technologies, ISD, and learning environments: Critical perspectives. *Educational Technology Research and Development*, 40 (1), 49-63.
- Hasselhorn, M. (1992). Metakognition und Lernen. In G. Nold (Hrsg.), *Lernbedingungen und Lernstrategien. Welche Rolle spielen kognitive Verstehensstrukturen?* (S. 35-63). Tübingen: Narr.
- Hermans, H., Petermann, F. & Zielinski, W. (1978). *LeistungsMotivationsTest*. Amsterdam: Swets en Zeitlinger.
- Issing, L.J. (1994). Von der Mediendidaktik zur Multimedia-Didaktik. *Unterrichtswissenschaft*, 22 (3), 267-284.
- Järvelä, S. (1995). The cognitive apprenticeship model in a technologically rich learning environment: Interpreting the learning interaction. *Learning and Instruction*, 5 (3), 237-259.
- Jonassen, D.H., Hannum, W.H., & Tessmer, M. (1989). *Handbook of task analysis procedures*. New York: Praeger.
- Käppler; Pitrella (1988). Identification and Evaluation of Scale Design Principles
- Lajoie, S.P., & Lesgold, A. (1989). Apprenticeship training in the workplace: Computer-coached practice environment as a new form of apprenticeship. *Machine-Mediated Learning*, 3, 7-28.
- Langner, T. (1991). Analyse von Einflußfaktoren beim rechnergestützten Konstruieren. In Beitz, W. (Hrsg.), *Schriftenreihe Konstruktionstechnik* (Bd. 20). Berlin: TU Berlin.
- Litchfield, B.C., & Mattson, S.A. (1989). The interactive media science project: An inquiry-based multimedia science curriculum. *Journal of Computers in Mathematics and Science Teaching*, 9 (1), 37-43.
- Niegemann, H.M. (1995). Zum Einfluß von „modelling“ in einer computergestützten Lernumgebung: Quasi-experimentelle Untersuchung zur Instruktionsdesign-Theorie. *Unterrichtswissenschaft*, 23 (1), 75-87.
- Norman, D.A. (1978). Notes toward a theory of complex learning. In A.M. Lesgold, J.W. Pellegrino, S.D. Fokkema, & R. Glaser (Eds.), *Cognitive psychology and instruction* (pp. 39-48). New York: Plenum Press.
- Pfendler, C. (1993). Vergleich der Zwei-Ebenen-Intensitätsskala und des Task Load Index bei der Beanspruchungsbewertung während Lernvorgängen. *Zeitschrift für Arbeitswissenschaft* 47 (19 NF), 26-33.
- Rieber, L.P. (1996). Seriously considering play: Designing interactive learning environments based on the blending of microworlds, simulations, and games. *Educational Technology Research and Development*, 44 (2), 43-58.
- Riedel, K. (1973). *Lehrhilfen zum entdeckenden Lernen. Ein experimenteller Beitrag zur Denkerziehung*. Hannover: Schroedel.
- Ross, S.M. & Morrison, G.R. (1997). Measurement and evaluation approaches in instructional design: Historical roots and current perspectives. In R.D. Tennyson, F. Schott, N.M. Seel & S. Dijkstra (Eds.), *Instructional design: International perspectives*. Vol. 1: Theory, research and models (pp. 327-351). Mahwah, N.J.: Erlbaum.
- Samuelson, P. (1983). *Economics from the heart: A Samuelson sampler*. New York: Harcourt & Brace.
- Scandura, J.M. (1995). Theoretical foundations of instruction: Past, present, and future. *Journal of Structural Learning and Intelligent Systems*, 12 (3), 231-342.

- Schank, R.C., Fano, A., Bell, B., & Jona, M. (1993/94). The design of goal-based scenarios. *The Journal of the Learning Sciences*, 3 (4), 305-345.
- Seel, N.M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe.
- Seel, N.M. (1997). Lernumgebungen. Theoretische Grundlagen und Anwendungen. In J.P. Pahl (Hrsg.), *Lern- und Arbeitsumgebungen zur Instandhaltungsausbildung* (S. 11-46). Seelze-Velber: Kallmeyer.
- Seel, N.M., & Dörr, G. (1997). Die didaktische Gestaltung multimedialer Lernumgebungen. In H.F. Friedrich, G. Eigler, H. Mandl, W. Schnotz, F. Schott, & N.M. Seel (Hrsg.), *Multimediale Lernumgebungen in der betrieblichen Weiterbildung. Gestaltung, Lernstrategien und Qualitätssicherung* (S. 73-163). Neuwied: Luchterhand.
- Seel, N.M., Al-Diban, S., Held, S., & Hess, C. (1998). Didaktisches Design multimedialer Lernumgebungen. Theoretische Positionen, Gestaltungsprinzipien, empirische Befunde. In G. Dörr & K.L. Jüngst (Hrsg.), *Lernen mit Medien. Ergebnisse und Perspektiven zu medial vermittelten Lehr- und Lernprozessen* (S. 87-119). Weinheim: Juventa.
- Stolurow, L.M. (1973). Lernumwelten oder Gelegenheiten zum Nachdenken. In W. Edelstein, & D. Hopf (Hrsg.), *Bedingungen des Bildungsprozesses. Psychologische und pädagogische Forschungen zum Lehren und Lernen in der Schule* (S. 351-398). Stuttgart: Klett.
- Tennyson, R.D., Schott, F., Seel, N.M. & Dijkstra, S. (Eds.) (1997). *Instructional design: International perspectives. Vol. 1: Theory, research and models*. Mahwah, N.J.: Erlbaum.
- Volet, S.E. (1991). Modelling and coaching of relevant, metacognitive strategies for enhancing university students' learning. *Learning and Instruction*, 1 (4), 319-336.

Anschrift der Autoren:

Albert-Ludwigs-Universität Freiburg i.Br.  
Seminar für Philosophie und Erziehungswissenschaft  
Dipl.-Psych. Sabine Al-Diban und Prof. Dr. Norbert M. Seel  
Werthmannplatz  
79098 Freiburg  
email: seel@ezw.uni-freiburg.de